

Background Document for the Guide for P2P impact assessment

Organisation name of lead contractor for this document:

The University of Manchester (UNIMAN), United Kingdom

Key authors: Effie Amanatidou, Paul Cunningham, Deborah Cox

Contributors: Angus Hunter, Optimat, United Kingdom

Michael Dinges, Anja Köngeter, Susanne Meyer, AIT, Austria

Project no.	645782
Project acronym:	ERA-LEARN 2020
Project full title:	Strengthening joint programming in Europe
Funding scheme:	Coordination and support action
Start date of project:	1 January 2015
Duration:	36 months

Deliverable D 3.4b

Background Document on P2P evaluation / impact assessment

Due date of deliverable:	(M18)
Actual submission date:	(M18) Revised version (M24)
Dissemination Level:	PU
Organisation name of lead contractor for this deliverable:	The University of Manchester (UNIMAN), United Kingdom
Key authors:	Effie Amanatidou, Paul Cunningham, Deborah Cox
Contributors:	Angus Hunter, Optimat, United Kingdom Michael Dinges, Anja Köngeter, Susanne Meyer, AIT, Austria

Introduction

The ERA-LEARN 2020 project dedicates a specific work-package (WP3) to monitoring and impact assessment, which aims to implement a more integrated and systematic framework for monitoring and assessing the impact of public-to-public (P2P) networks and associated co-funded projects. The aim is to understand how P2Ps are to be evaluated as well as the value and usefulness of impact assessment of P2Ps including ERA-NETs, ERA-NET Cofund, JPIs and Article 185 initiatives.

This document is the Background Document that accompanies the ‘Short Guide on P2P evaluation / impact assessment’ (Deliverable D3.4a). As such it provides additional information on a number of the concepts used in the Guide as well as examples from P2Ps relevant work. Naturally, the Background Document is a living document that can and should be updated in the course of ERA-LEARN 2020 reflecting on the rising needs for evaluation / impact assessment of P2Ps. It is also important to note that the term ‘impact assessment’ in the Guide and this Background Document is seen in the wider sense addressing also evaluation of an intervention’s effectiveness, efficiency, etc. as well as its outputs and outcomes. Thus, the terms ‘evaluation’ and ‘impact assessment’ are used interchangeably.

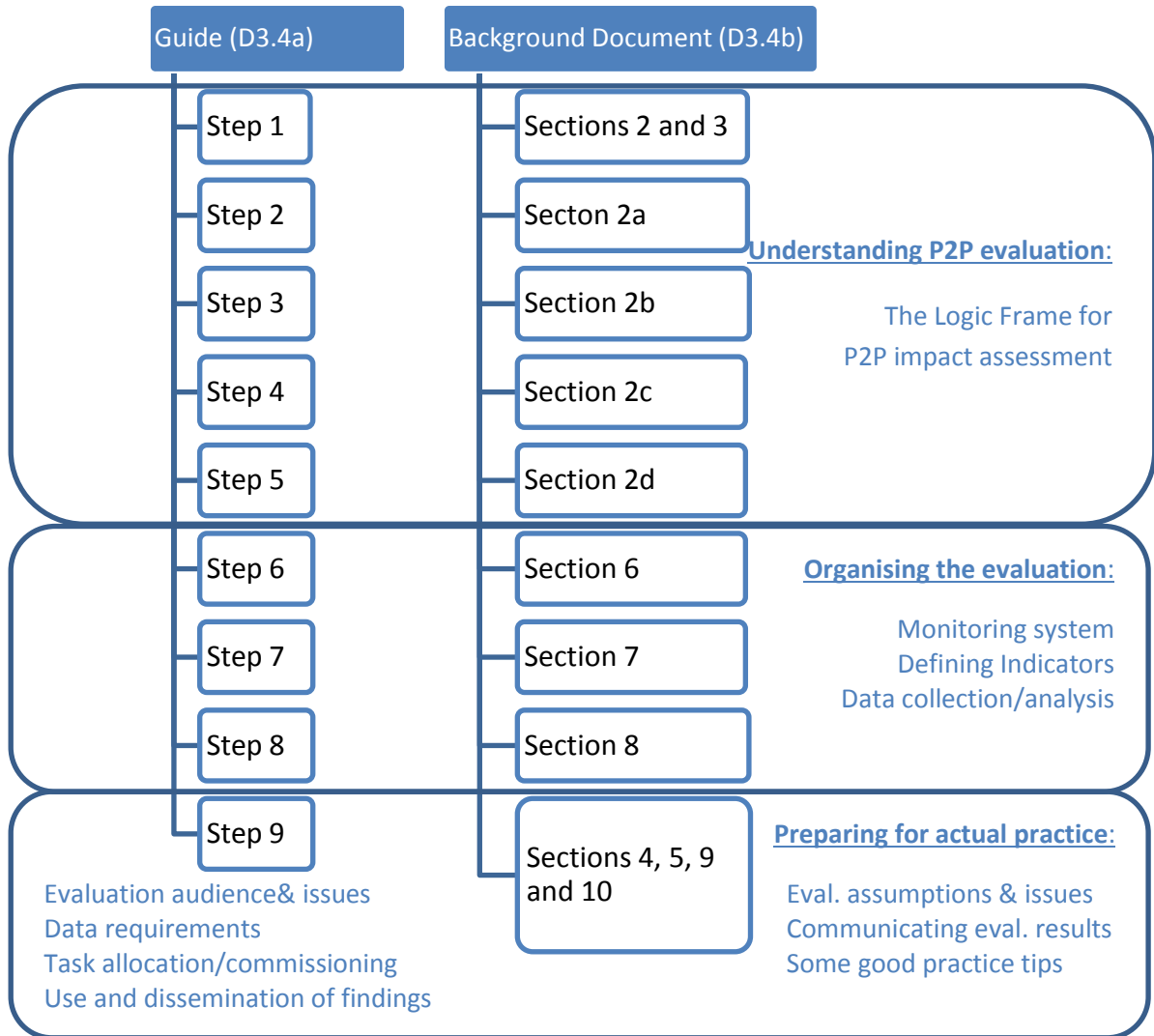
The report draws upon

- a broad literature review on the state-of-the-art practices in evaluation and impact assessment of research programmes, as well as existing guide books for impact assessment of programmes, initiatives, and research networks;
- the work carried out already by certain P2Ps that have developed their evaluation frameworks and already carried out evaluation/assessment exercises; and
- the results of the discussions during the ERA-LEARN 2020 workshop on evaluation of P2Ps that was carried out in Brussels, 18th May 2016.

Both the Guide and the Background Document are addressed to P2P management actors responsible for evaluation tasks in their networks and aim to

- establish a common evaluation language / understanding of terms and concepts
- Improve understanding of the value and usefulness of impact assessment activities and results
- address the needs of the different stakeholders, i.e. higher-level policy-makers and public officials at European, national and regional levels as well as P2P network coordinators and participants, in identifying and demonstrating P2P impacts;
- help the P2P community conduct IA exercises alone or assign it to external evaluators
- provide advice and recommendations on how to use impact assessment results at both the strategic/programme and operational levels.

The correspondence between the structures of the two documents ‘Guide for P2P evaluation / impact assessment’ (D3.4a) and the Background Document (D 3.4b) is shown overleaf.



Contents

Introduction	3
1. Special features of P2Ps	6
2. Basic elements of a Logical Framework for P2P impact assessment.....	7
a. The Challenges or problems or needs targeted by a P2P	9
b. The Objectives of P2Ps.....	9
c. The inputs and activities of P2Ps	13
d. Outputs, outcomes and impacts of P2Ps.....	14
Impacts at project level of a P2P.....	15
Impacts at P2P network level.....	16
Periodicity of impacts	17
3. A suggested Logical Framework for P2P evaluation	17
4. The assumptions underlying the design of a P2P	21
5. Evaluation issues in P2P evaluation	24
Setting the appropriate time for an evaluation.....	27
6. The importance of a monitoring system.....	29
Project level information collection.....	30
Programme / instrument level information collection.....	31
7. Defining output, outcome and impact indicators.....	34
8. Data collection and analysis methods	40
9. Communicating evaluation / impact assessment results	42
10. Some good practice tips.....	44
11. Sources / suggestions for further reading:	46

1. Special features of P2Ps

Public-to-Public partnerships (P2Ps) involve ERA-NET, or ERA-NET Plus or ERA-NET Cofund Actions, Art 185 Initiatives and Joint Programming Initiatives. They are partnerships among public agencies and/or ministries responsible for research and innovation policies and/or programmes across different countries with the overall aim to pool resources and capacities in pursuing jointly agreed targets for research and innovation.

Joint Programming Initiatives (JPIs) are partnerships formed by interested Member States (at the highest political level, i.e. Ministries) with the purpose of jointly implementing a Strategic Research and Innovation Agenda (SRIA), which is agreed among Member States to address major societal challenges. Member States commit to JPIs on a variable geometry basis.

Article 185 Initiatives engage interested Member States in a collaborative effort to implement a jointly developed research programme that is co-funded by national/regional budgets/programmes as well as the European Commission. The participating EU Member States integrate (in terms of scientific, managerial and financial aspects) rather than simply coordinate their research efforts by defining and committing themselves to this joint research programme.

The ERA-NET scheme aims at developing and strengthening the coordination of national and regional research programmes through the collaboration of national and regional authorities, represented by so-called 'programme owners' and/or 'programme managers'. ERA-NET actions have had different roles during the Framework Programmes: 'ERA-NET actions' in FP6 provided support to coordinate national/regional activities by developing joint calls for trans-national proposals. 'ERA-NET Plus actions' in FP7 provided - in a limited number of cases with high European added value - additional EU financial support to top-up the research funding of a single joint call for proposals between national and/or regional programmes. The ERA-NET Cofund instrument under Horizon 2020 merges the former ERA-NET and ERA-NET Plus into a single instrument with the central and compulsory element of implementing one substantial call with top-up funding from the Commission but also allowing for additional activities.

Despite the differences in technicalities and level of reference (ministry level, agency level, etc.) of the three different types of P2Ps, they are all structures that are formed for a specific time, and might be slightly or significantly changed over time in terms of membership and thus capacities and committed resources. In other words, they are networks of agencies and/or ministries that join forces to pursue commonly agreed targets. *A network is a decentralized member-driven platform of relationships that evolves its capabilities and underlying structure of connectivity.*¹

¹ Network Impact and Center for Evaluation Innovation. 2014. Part 1 of a Guide to Network Evaluation Framing Paper: The State of Network Evaluation, July 2014.

As networks

- they have numerous players, some of whom may leave after a given period of time (e.g. after a Cofund Action ends) while others may join in when a follow-up action is planned;
- not all members need to join and commit resources to realise the planned, joint activities; thus there are diverse types and level of membership and engagement;
- their success depends on the degree to which the network establishes connections among its members building trust, and long-term commitment.

2. Basic elements of a Logical Framework for P2P impact assessment

A useful first step in conducting an evaluation is to construct what is termed a 'logical framework' or 'log-frame' for the P2P. This provides an overall roadmap in which the rationale, objectives, activities, results, outcomes and impacts can be situated in a logical and interconnected context.

A Logical Framework or Logic Model/Frame is “a management tool used to improve the design of interventions, most often at the project level. It involves identifying strategic elements (inputs, outputs, outcomes, impact) and their causal relationships, indicators, and the assumptions or risks that may influence success and failure.”²

Generally, a logic model will identify the following elements of a policy intervention:

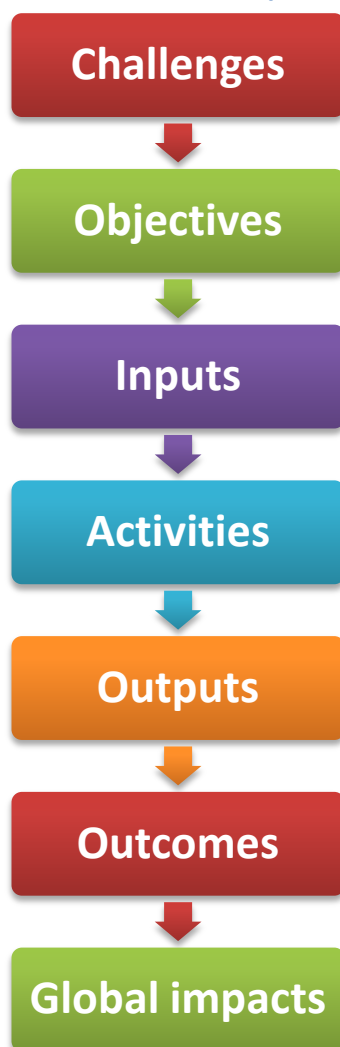
- the issues being addressed and the context within which the policy takes place;
- the inputs, i.e. the resources (money, time, people, skills) being invested;
- the activities which need to be undertaken to achieve the policy objectives;
- the initial outputs of the policy;
- the outcomes (i.e. short and medium-term results);
- the anticipated impacts (i.e. long-term results); and
- the assumptions made about how these elements link together which will enable the programme to successfully progress from one element to the next

Source: The Magenta Book, <https://www.gov.uk/government/publications/the-magenta-book>

The basic elements of such a frame are illustrated as follows. When preparing for evaluating or assessing the impact of a P2P we need to clearly define and identify each of these elements.

² OECD (2009) Glossary of Key Terms in Evaluation and Results Based Management. <http://www.oecd.org/development/peer-reviews/2754804.pdf>

Figure 1: Basic elements of Logical Framework for P2P impact assessment



In the following sections (2a – 2d) each of the framework components is presented in detail and discussed in the context of P2Ps.

a. The Challenges or problems or needs targeted by a P2P

The **challenge or problems or needs** that the P2P tries to deal with may be societal challenges that the Member States agree that are important to find solutions for based on research and innovation. Yet, they may also refer to other types of issues like increasing industrial competitiveness in certain sectors.

For instance, the Joint Programming Initiative JPND deals with the following challenges.

- The ageing of the European population
- The number of European citizens suffering from neurodegenerative diseases
- The rising costs of healthcare
- The debilitating and largely untreatable character of disorders that are linked with age
- A need to improve understanding of neurodegenerative diseases and to provide new approaches for prevention, diagnoses and treatment
- A need to effectively provide healthcare, social care and support to optimise quality of life at all stages of the illness

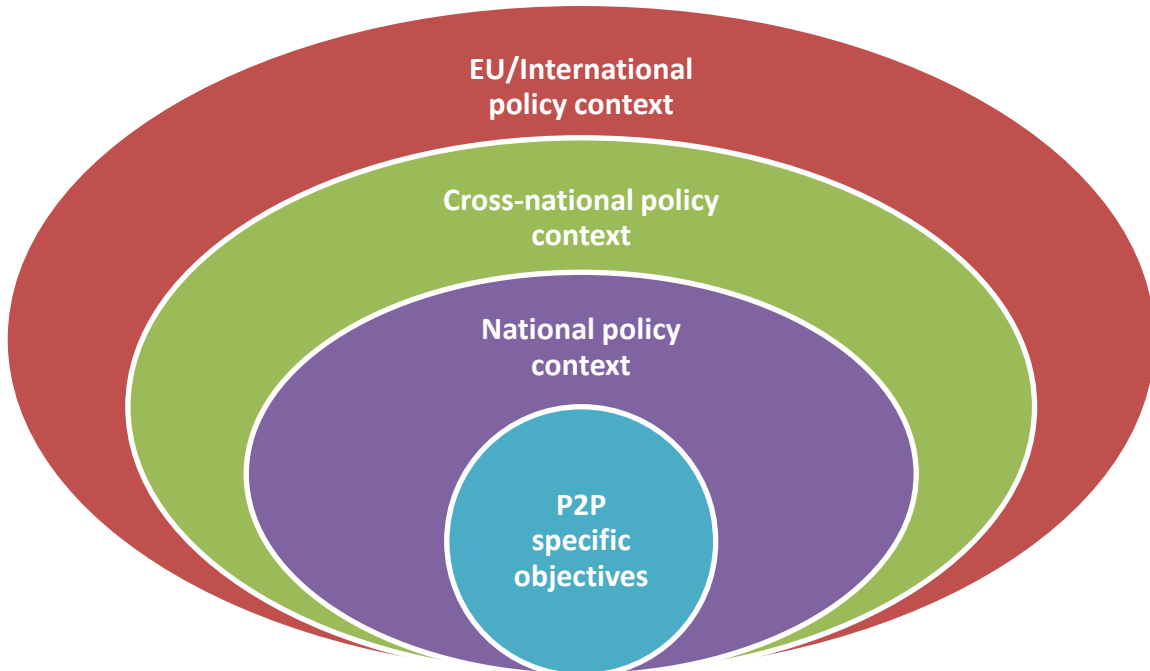
JPI MYBL deals with the challenge and demographic change and its implications in economies and societies of today and particularly on healthcare and long-term care, social and welfare systems, the way we work and learn as well as on infrastructures, cities and transport systems. The relevant site (<http://www.jp-demographic.eu/about-us/background-and-goals/>) notes “All approaches to tackle the implications of demographic change and to provide guidance in shaping the transition towards a human, socially inclusive society need an active participation of politics, economy, and research.”

The ERA-NET BiodERsA deals with the challenge of preserving biodiversity. “Biodiversity is Earth’s living heritage and is essential for the delivery of ecosystem services. It is as much an environmental issue as an economic, food-security, energy-security and political one. With adequate support and a solid research basis, biodiversity and ecosystem services can become a toolbox for the resilience of human societies facing a changing, unsettled world. Biodiversity questions cross both borders (e.g. alien invasive species) and disciplines (e.g. valuation of ecosystem services). They can thus be tackled more effectively by collaboration between multi-national teams.” (<http://www.biodiversa.org/pdf/brochurebiodiversa.pdf>).

b. The Objectives of P2Ps

The **objectives** a P2P was designed to achieve – these objectives are usually implied under a broader rationale for taking action as the example of JPND shows above. As the P2P is not isolated of the wider policy context at both national and European level (if not international), and in most cases it aims to contribute to achieving wider policy aims, its objectives need to be placed as part of an objectives’ hierarchy considering the broader policy context surrounding the intervention.

Figure 1: Positioning a P2P within the wider policy contexts



Naturally, not all policy contexts need to be addressed in an objectives' hierarchy unless they are directly relevant to the nature and content of the P2P. The objectives' hierarchy would then consist of the following types of objectives.

Definitions

Operational objectives provide a basis for assessing a P2P in relation to its outputs. The latter can be defined as what is directly produced / supplied through the activities and actions carried out during the implementation process.

Specific objectives provide a basis for assessing a P2P in relation to the short-term results that occur at the level of direct beneficiaries/recipients of assistance.

Intermediate objectives provide a basis for assessing a P2P in relation to its short to medium-term effects (or intermediate impacts) on both the direct and indirect beneficiaries/recipients of assistance.

Global objectives provide a basis for assessing a P2P in relation to longer term and more diffuse effects (or global impacts).³

Source: Adjusted from European Commission 2004

³ Outputs, outcomes and impacts are further discussed in the relevant section below.

As an example the objectives hierarchy of the EMRP/EMPIR Art 185 Initiative is shown below. The EMRP/EMPIR’s specific objectives are reflecting the overall objectives of the whole intervention, i.e. the EMRP/EMPIR programme, while the operational objectives reflect the outputs of each activity planned and the global objectives are those of the wider policy within which EMRP/EMPIR is placed, i.e. H2020 and the ERA objectives.

Figure 2: EMRP/EMPIR Objectives Hierarchy



Source: Paula Knee, 2016, Article 185 Impact Assessment EMRP /EMPIR. Presentation at the ERA-LEARN 2020 Workshop on evaluation and impact assessment of P2Ps, Brussels 18 May 2016⁴.

Another example of objectives’ hierarchy comes from the JPND case where the objectives are not grouped in terms of specificity and timing but in terms of their relation to policy, science and society. These objectives can also be translated into more specific, operational intermediate and global objectives.

⁴ <https://www.era-learn.eu/events/era-learn-2020-workshop-on-evaluation-and-impact-assessment-of-p2ps>

Table 1: Translating JPND objectives into an Objectives' Hierarchy

<i>Policy-related objectives</i>	<i>Type of objective</i>
Identification of common goals that would benefit from joint actions	Short-term, specific/operational
Alignment of national research programmes	Longer-term
Creation of critical mass of research capacity	Medium-term, intermediate
Implementation of experience into evidence-based policies and best practices	Longer-term, global
<i>Science-related objectives</i>	
SRA-related objectives	Short-, medium-, longer-term
Stimulation of education and training of research professionals	Medium-term, intermediate
<i>Society-related objectives</i>	
Destigmatisation of patients	Medium-term, intermediate
Raising awareness about the importance of neurodegenerative research	Short-term, specific/operational

Another example comes from MNT ERA-NET based on reading the available documents.

Figure 3: Objectives Hierarchy of MNT ERA-NET (<http://www.mnt-era.net/MNT/>)



c. The inputs and activities of P2Ps

Inputs are the means used to support activities and action and to produce outputs. Inputs include budgetary costs (financial, administrative and human resources), but also costs for the beneficiaries or target population (co-financing, compliance costs stemming from administrative burden) and costs for third parties (Member States, intermediary organisations, etc.). Inputs need to be documented and monitored as one of the main evaluation issues that is usually examined is efficiency, i.e. the extent to which the desired effects are achieved at a reasonable cost.

At the same time, the management and governance structures, and processes governing the operation of the P2P, i.e. how the P2P is set to operate may also be regarded as inputs. These elements are of major importance when assessing another evaluation issue, that of network health and connectivity.⁵

The **activities** through the implementation of which we expect the desired impacts to occur and the set objectives to be achieved. P2Ps broadly aim at the coordination of national/regional research and innovation activities and the collaboration between national/regional funding organisations. Apart from implementing transnational calls there is a range of other possible joint activities that have already been implemented by P2P networks⁶.

- Implementing transnational calls
- Additional joint calls
- Mapping national/trans-national activities
- Foresight and common vision
- Strategic Research Agenda / Implementation Plan
- Knowledge sharing amongst researchers
- Mobility and training of researchers / activities promoting early career scientists and young researchers
- Research infrastructures
- Stakeholder involvement
- Dissemination / Up-Take of research results
- Widening participation (activities related to extending cooperation with less active EU countries)
- Internationalisation (activities related to extending cooperation to third / non-EU countries)
- Monitoring and evaluation/assessment activities both in relation to the network itself or the co-funded projects
- Other activities that support the alignment of national programmes
- Other joint activities

⁵ Evaluation issues are presented in detail in section 5 below.

⁶ Each of these is further described in <https://www.era-learn.eu/joint-activities>.

d. Outputs, outcomes and impacts of P2Ps

The activities supported by the inputs invested lead to certain outputs; these are usually tangible **outputs** such as projects supported under joint calls, joint strategy documents stemming from strategy building activities, training modules of students/researchers trained, databases with mapped national/regional programmes, new partners from different Member States brought together, etc.

The activities carried out cause interactions among individuals, and organisations, blending of minds, creation of links, etc. This does not stop with the production of the agreed outputs. In fact, delivery of outputs reinforces the relations and interactions among the partners or beneficiaries of a P2P; thus it increases collaboration. The growing collaboration among the P2P partners as well as the engagement of beneficiaries leads to **medium term impacts (or otherwise outcomes)** on the P2P target group(s). These impacts may be directly associated with the outputs such as improved skills and capacity building from training activities, research results from the projects supported, new collaborations among P2P beneficiaries, etc. They can however include impacts that are not directly associated with the outputs produced but have more to do with the interactions and increasing collaboration among P2P partners and beneficiaries, i.e. process impacts, for instance increased trust and improved collaboration among partners, increased awareness of a policy area at national or cross-national level, etc.

The emergence of short to medium-term impacts may be strengthened even further by a favourable wider context. They can translate, either intentionally or even unintentionally, to long-term impacts on target groups as well as society and economy at large. These are called **global impacts**. Given that it takes time for such type of impacts to occur, attribution of these to the specific policy intervention is rather difficult. In other words, the more time passes from the completion of the intervention the less it is possible to attribute any effects to the specific intervention.

Comparing the three types, outputs are items directly produced by certain activities (e.g. workshop reports, SRAs, databases of programmes, etc.) and they are produced within the short-term. Intermediate impacts are rather medium-term and may refer to both direct and indirect beneficiaries while global impacts are longer-term and refer to the wider environment surrounding the policy intervention.

Adding to this, there are different types of impacts depending on their content such as scientific impacts, innovation-related, societal, cultural, environmental, etc. While these impacts come from the conduct of research i.e. at the project level of P2Ps, there are others that relate to the networking element of a P2P. These refer to connectivity impacts of P2P members, structural impacts, etc. as shown below. Naturally, it is not possible to search for all impact types at the same time. The types of impacts to be anticipated depend on the nature of the activity/ies examined, the timing of the evaluation and the stage of development of the P2P. The two lists below are illustrative of all possible different types of impacts at network and project level. In a P2P evaluation certain types can be selected instead of others. The choice of impacts to be retained for deeper assessment should be clearly justified.

Impacts at project level of a P2P⁷

Science impacts: research results have an effect on the subsequent progress and contribution to the body of knowledge. They affect the formation and development of disciplines as well as training and can also affect the development of a research field itself, generating interdisciplinary and international projects.

Innovation impacts: product, process and service innovations as well as know-how partly result from research activities. There are few indicators for assessing this dimension, other than patents, which have generated some debate regarding their utility.

Economic impacts: these refer to the impact on an organisation's budgetary situation, operating costs, revenues, profits, the sale price of products; on the sources of finance, investments and production activities; and on the development of new markets. At the aggregate level, they can also refer to economic returns, either through growth or increased productivity, of a given geographical unit.

Cultural impacts: these relate to an individual's knowledge and understanding of ideas and reality, as well as intellectual and practical skills, attitudes, interests, values and beliefs.

Societal impacts: research affects the welfare, behaviour, practices and activities of people and groups, including their well-being and quality of life.

Policy impacts: research influences how policy makers and policies act. It can provide evidence that influences policy decisions and can enhance citizens' participation in scientific and technological decisions.

Organisational impacts: these refer to the effects on the activities of institutions and organisations: planning, organisation of work, administration, human resources, etc.

Health impacts: these relate to impacts on public health, e.g. life expectancy, prevention of illnesses, quality of life, and the health-care system.

Environmental impacts: these concern management of the environment, notably natural resources and environmental pollution, as well as the impacts of research on climate and meteorology.

Symbolic impacts: these are the gains in areas such as credibility due to undertaking R&D or linked to universities or research institutions that offer gains in terms of potential clients, etc.

Training impacts: these are impacts of research on curricula, pedagogical tools, qualifications, entry into the workforce, etc.

⁷ Source: adjusted from CIA4OPM, 2011

Impacts at P2P network level⁸

Enduring connectivity relates to the on-going communication between the relevant actors and to the follow on collaborations that continue after the initial activity has been completed. This is connectivity that lasts beyond the first funded relationship. This type of impact refers to both P2P member organisations as well as their final beneficiaries. Retained collaboration through new projects or networks can be one example.

Capacity building refers to the development/improvement of capabilities and skills in the P2P member organisations. The areas of skill development may include international programme/programme management, strategic thinking, international collaboration, international team coordination, etc.

Attitudinal/cultural change relates to knowledge exchange and includes elements such as improved reciprocal understanding and willingness to work together among P2P current and potential partners. It may also refer to changes in research organisation such as adopting multidisciplinary approaches in research.

Conceptual impact refers to the impact on the knowledge, understanding and attitudes of policy-makers. In this category of impact we identify examples of changed thinking amongst policy makers, influences on policy issues and increased awareness in the policy world due to participation in a P2P. This type of impact at P2P network level can be paralleled to the policy impacts that may result from P2P funded projects. Policy impacts can influence national as well as European or international policies or strategies.

Structural impacts relate to changes in institutions and structures in the national or European research landscape due to changed thinking amongst policy makers and influences on policy issues stemming from the acquired knowledge from participating in a P2P. The set-up of specific formal or informal structures in order to improve coordination at national level is an example of this type.

Economic and symbolic impacts may also occur at the network level referring for instance to increased national investment in a specific area through P2Ps or reputation benefits due to increased international profiles of P2P partners

⁸ Adjusted from Meaghar's impact framework
http://www.ruru.ac.uk/pdf/oct2014/Laura_Meagher_presentation.pdf.

Periodicity of impacts

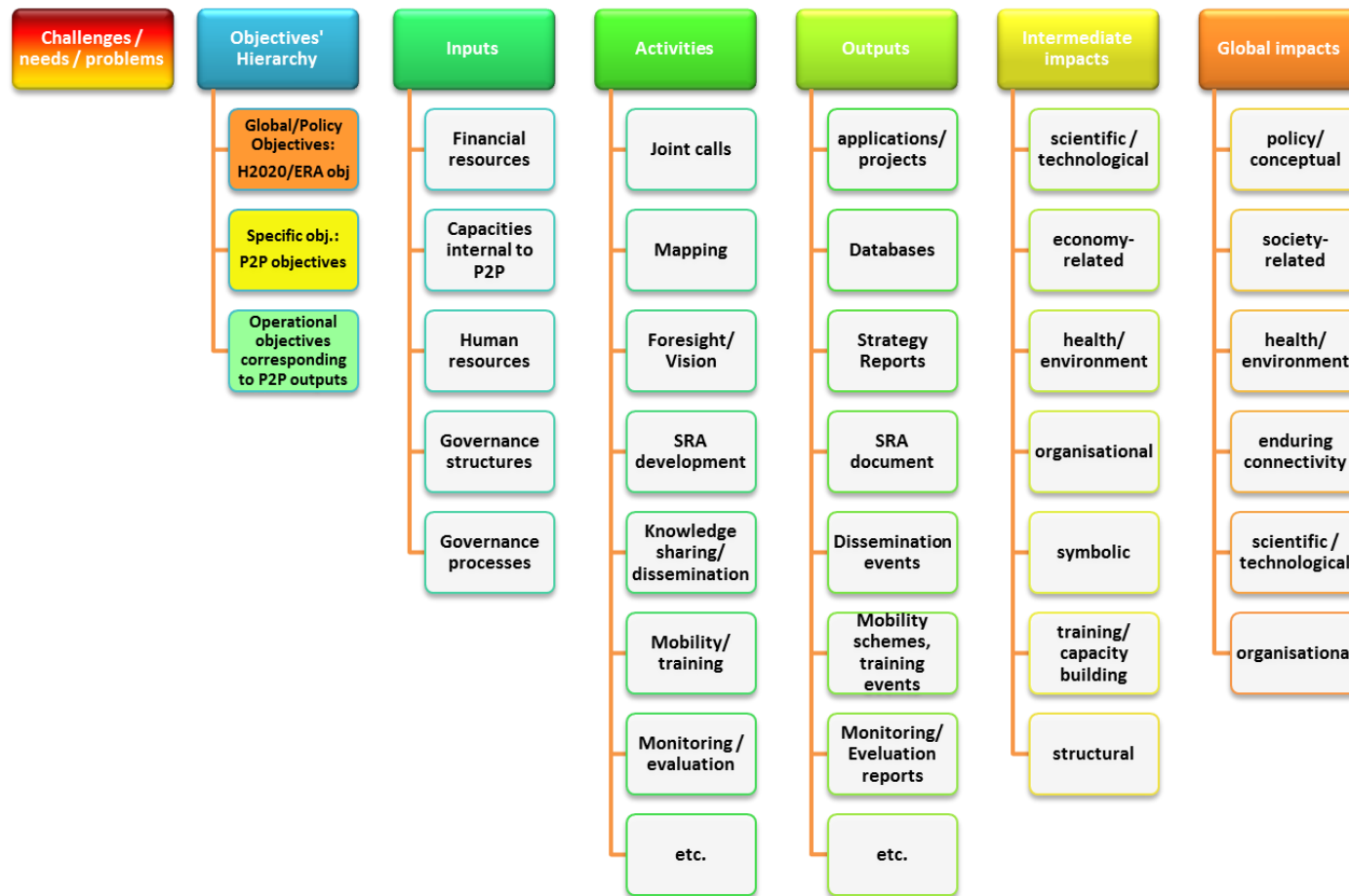
Impacts can also vary in terms of when they manifest. For instance, at network level, capacity building impacts and connectivity were the first types of impacts that were appreciated by participants in P2P supported projects based on JPI interviews conducted under ERA-LEARN 2020 in 2015. There were also signs of certain policy/conceptual types of impacts and some cultural impacts. The latter would typically be of medium to long-term nature as it takes time to change mindsets or long-established policies or to develop new policies. Enduring connectivity would be the type of impact to examine in the long-term.

Scientific and technological impacts are rather short- to medium term, being more directly associated with the results of the supported projects. The same goes for organisational impacts. Yet, in both cases some initial impacts may emerge in the short-term and, depending on the conditions in place, they may evolve into further impacts in the medium to long-term. For instance, the scientific results of a project may lead to a specific development in the future if taken up after the end of a project. Participation in a P2P network may require an organisation to change some rules in programme management, which in the medium-term may lead to internal changes in the whole management practices applied in the organisation. In a similar way, economic, health and environmental impacts may exhibit both a short and long-term nature.

3. A suggested Logical Framework for P2P evaluation

Based on the above, an illustrative Logic Frame for evaluation and impact assessment of P2Ps would look like this (Figure 6).

Figure 4: An illustrative Logic Frame for evaluation and impact assessment of P2Ps



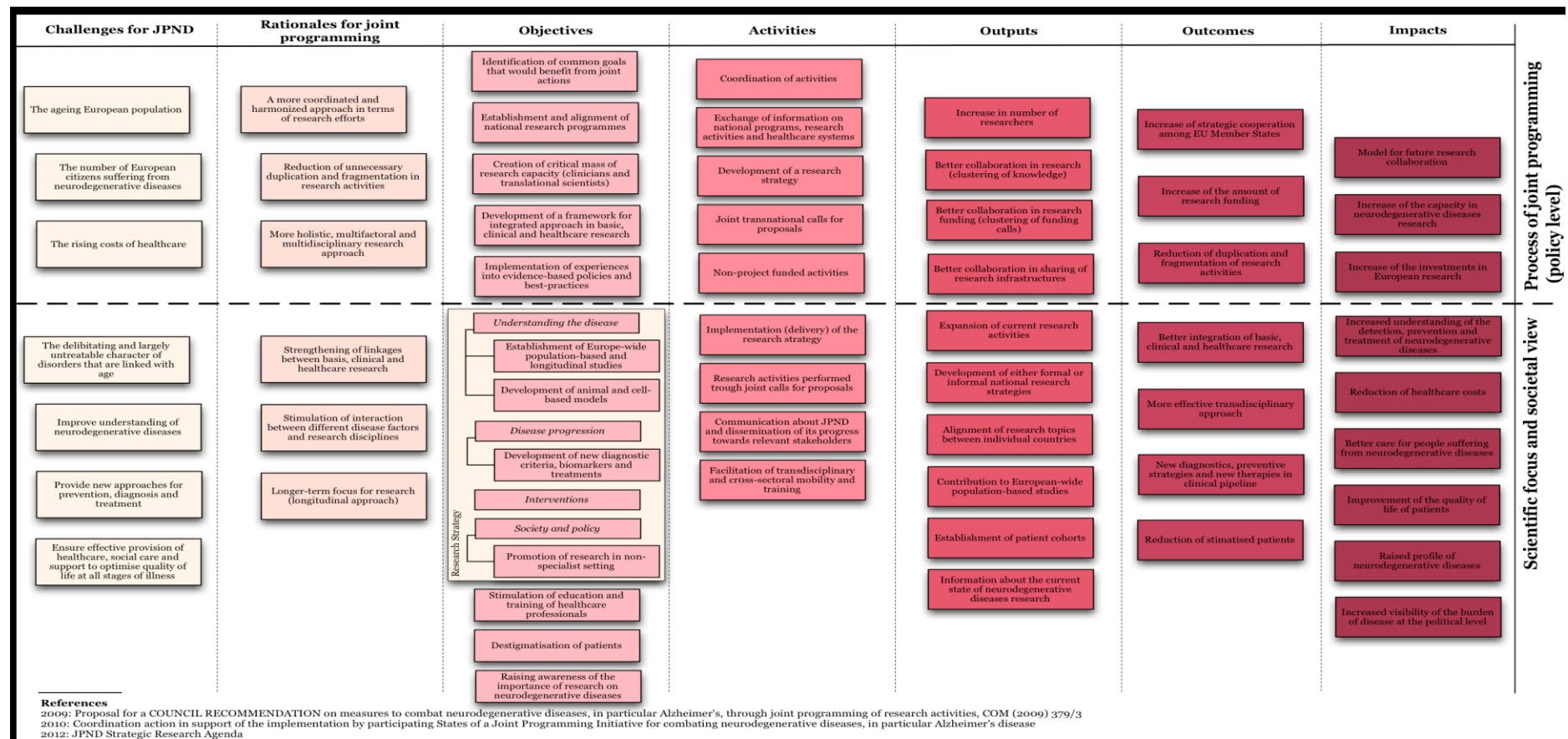
The respective example of a Logic Frame for the evaluation of the EMPIR/AMRP Art. 185 looks as follows.

Inputs / activities	Outputs	Outcomes	Impact
<p>ADMINISTRATIVE National & EU financial resources Governance structures & processes Management & administration structures & processes Develop Strategic Research Agenda Human resources Training / capacity building</p> <p>PROJECTS Design, select & conduct collaborative metrology research</p>	<p>INTEGRATION 50% of national metrology research investments in Europe coordinated via EMRP/ EMPIR Strategic Research Agenda Effective governance, management and administration deployed: Centralised processes for project selection, monitoring, payments Common rules & processes for grant agreements, costings, etc. Improved capacity of developing NMIs</p> <p>PROJECTS Scientific papers Contribution to standards committees, working groups Training Patents Communication & dissemination</p>	<p>INTEGRATION Progress towards financial / management/ scientific integration</p> <p>PROJECTS Adoption of technical outputs by measurement users to support product and process improvements & innovation</p> <p>New or revised standards (ISO, CEN, etc) in use</p>	<p>Integrated European metrology landscape</p> <p>Increased business growth and productivity</p> <p>Solutions to the challenges of climate change, improved healthcare and security</p>

Source: https://www.era-learn.eu/events/annual-joint-programming-conference-2016/ERA_Learn_Plenary_2_PaulaKneeSchreibgeschlzt.pdf.

Another example of a Logic Frame comes from JPND. Interestingly in this case a distinction is made between aspects related to the process of joint programming (the upper part of the diagram) and the scientific focus and societal view of JPND (the lower part). This separation can also be paralleled to the network level and project level although some elements in the lower part of the figure would also be relevant at the network level as well.

Figure 5: JPND’s Logic Frame for evaluation



Source: Technopolis Group analysis, based on JPND documentation (2012).

4. The assumptions underlying the design of a P2P

In designing an intervention we usually make implicit or explicit **assumptions** about how the intervention would operate and how impacts would materialise. These assumptions essentially describe how we believe the different elements of the intervention (inputs, activities, outputs, etc.) would link together in reality in pursuing the set objectives. In other words, these are the assumptions that explain how the anticipated results and impacts are to be delivered by the intervention’s inputs and activities. For instance, when we design a training activity and link it with increased capacity skills within an organisation as an expected impact we assume that the person that will be trained will then share their knowledge with other colleagues, i.e. that there are procedures promoting knowledge transfer/sharing within the particular organisation, or that the person would be willing to do so on their own initiative.

These assumptions need to be clearly spelled out in a P2P evaluation framework. What is usually overseen, however, is the importance of examining why something was achieved apart from what has been achieved. This would enable lessons to be learnt for improving the design of future interventions. It is in the assumptions made that the reasons for success or failure of an intervention are usually hidden. The main assumptions that are made on how (by what activities, outcomes and impacts and by which interplays among them) the desired objectives are envisaged to be achieved comprise the **intervention logic** of the P2P.

Examples of assumptions that are relevant to P2Ps are provided below. Needless to say these need to be adjusted / revised in the case of a specific P2P evaluation or impact assessment exercise.

Assumptions in relation to inputs

- The governance structures and processes established are appropriate and adequate to build trust and trigger interest for participation at both the political and institution/organisation levels.
- There are sufficient resources (capacities, money, and infrastructure) secured to implement the planned activities.
- All members share a common purpose for the P2P and perceive its overall aims in the same way.
- The P2P members understand in a similar way the benefits of working together and jointly pursuing the set objectives. There is adequate political and financial commitment.

Assumptions in relation to activities and outputs

- There is shared understanding among the P2P members who are participating in the same activity.
- There are sufficient resources to ensure smooth implementation of all planned activities.
- The activities are designed in a way that will lead to the expected outputs.
- There is adequate interest and trust among P2P members to identify common areas of collaboration and implement coordination/collaboration work-plans effectively.

Assumptions in relation to activities and outputs vis-à-vis impacts

- If implemented successfully, the planned activities and outputs will lead to the envisaged intermediate and long-term impacts.
- The outputs will be of sufficiently high-quality to achieve the envisaged impacts

- Monitoring and evaluation will provide useful feedback to improve the design and operation of the P2P network at key points in time.
- There is adequate research capacity in the participating countries to respond to the joint call(s).
- Adequate efforts have been made to promote the joint call(s) within the participating countries.

Assumptions in relation to impacts vis-à-vis objectives

- The joint programming process enables effective coordination and cooperation among P2P members and contributes to (strategic and operational) alignment among P2P members.
- The necessary conditions are in place for the medium and long-term impacts to occur.
- The P2P impacts do not 'contradict' each other.

An example of assumptions comes from the evaluation framework developed by JPI AMR where the main assumption is that Joint Programming is about:

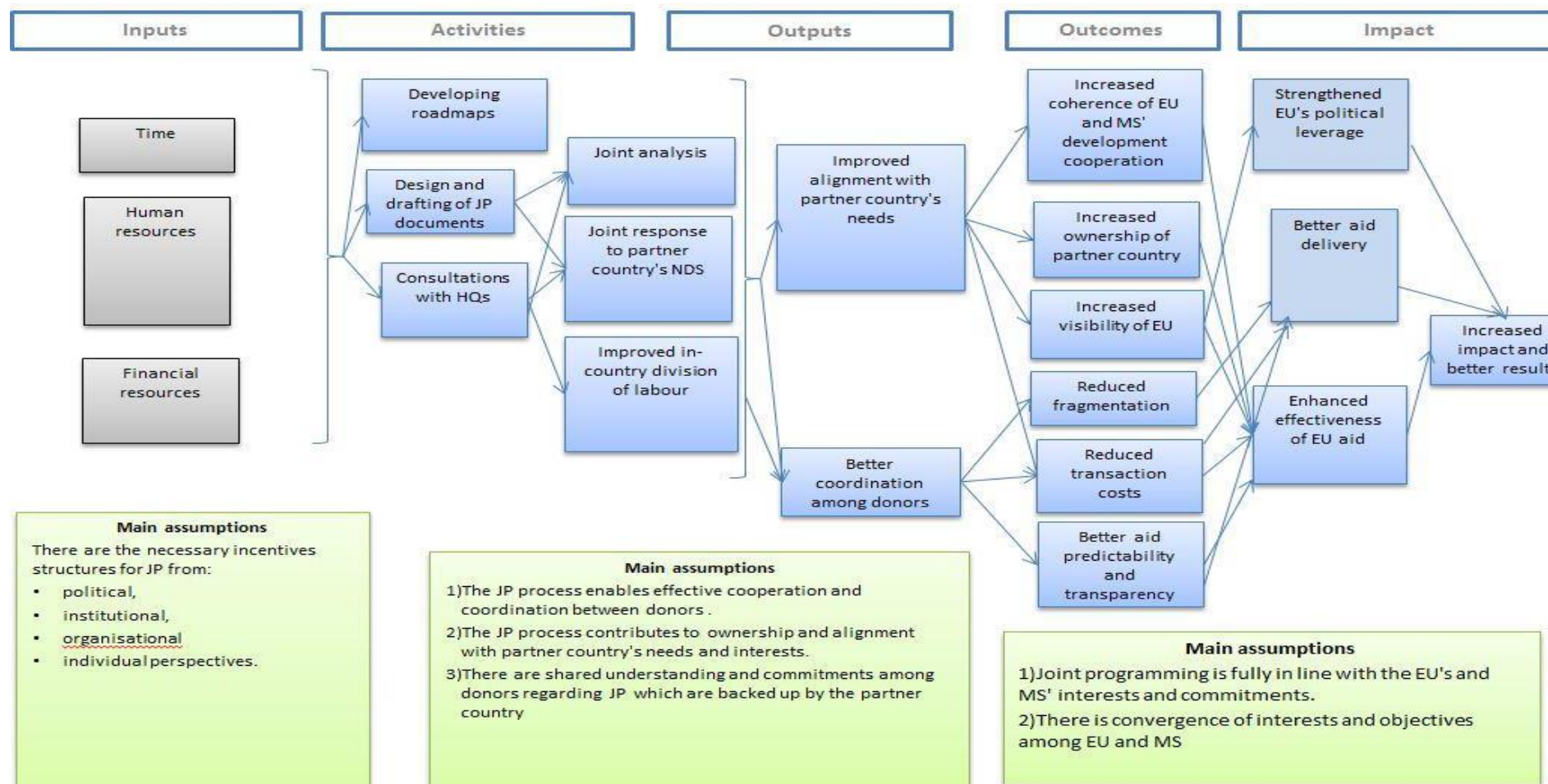
- Getting research decision makers to interact and collaborate towards a common goal – the societal challenge
- Getting researchers and resources to interact and collaborate towards a common goal – performing the best possible research in Europe to find better ways of addressing the societal challenge
- Facilitating the uptake of research outputs by those facing the challenge

Thus, it is argued that the performance and results of JPI AMR need to be assessed along these three dimensions: a) Governance of research policy making, b) Guiding research performance, and c) Addressing societal needs and innovation. This resembles the separation of the main objectives into policy-related, science-related and society-related as in the case of JPND.

Another example of the assumptions made about the various elements of a Logic Frame is given below (cf. Figure 7)⁹. It needs to be clarified that in DG DEVCO, joint programming has a different aim. It is seen as a process whereby the EU and MS (and other interested donors and partners) take strategic decisions based on a comprehensive view of donors' support to a given partner country and jointly plan development cooperation by the EU development partners working in a partner country. The overall aims of Joint Programming (for DG DEVCO) are: a) enhance the effectiveness and coherence of EU and MS aid, b) increase impacts and improve results, c) reduce fragmentation and increase transparency, predictability and accountability, and d) be open to all relevant development stakeholders.

⁹ Source: DG DEVCO, Evaluation of Joint Programming http://ec.europa.eu/smart-regulation/roadmaps/docs/2016_devco_004_evaluation_roadmap_jp_en.pdf

Figure 6: Logic Frame - An example of the main assumption underlying the intervention logic of JP by DG DEVCO



5. Evaluation issues in P2P evaluation

The above elements are the main components of an impact assessment framework. However, the framework is not complete yet. Whereas it shows which components of a policy intervention we need to identify and clarify in an evaluation exercise, it needs to be complemented with what exactly we will examine. This refers to the evaluation issues. **Evaluation issues** are essentially addressing the relations across the different elements of the framework and come from specific questions that we have in mind.

For instance the question ‘to what extent is an intervention relevant with respect to the needs, problems and issues identified in target groups?’ refers to the issue of Relevance. The question ‘to what extent do the effects induced by an intervention correspond with its objectives as they are outlined in the intervention strategy?’ refers to the issue of Effectiveness. The question ‘how economically have the resources used been converted into effects?’ refers to the issue of Efficiency. The question ‘how do the effects of an intervention compare with the wider needs of the target populations?’ refers to the issue of Utility. There are also other evaluation issues that can apply to P2Ps. A list of suggested issues to examine in a P2P evaluation is presented below.

Suggested evaluation issues for P2Ps and example questions

‘Relevance’ relates to the extent to which the P2P objectives are pertinent to the needs, problems and issues to be addressed.

Example question(s)

To what extent are the P2P objectives relevant with respect to the needs, problems and issues identified?

‘Coherence’ is the extent to which the intervention logic¹⁰ of a P2P is non-contradictory/the intervention logic does not contradict other interventions with similar objectives.

Example question(s)

To what extent is the intervention logic of the P2P compatible or in synergy or complementing other interventions with similar objectives?

‘Effectiveness’: the extent to which the set objectives and the intended results and impacts are achieved.

Example question(s)

To what extent do the effects induced by the P2P correspond with its objectives?

‘Efficiency’ refers to the extent to which the desired effects are achieved at a reasonable cost (in terms of resources consumed, such as time, financial inputs, etc.).

Example question(s)

How economically have the resources used been converted into effects?

¹⁰ The intervention logic is explained in the following section

'Utility' refers to the extent to which outcomes corresponded with the needs, problems and issues to be addressed.

Example question(s)

How do the effects of a P2P compare with the wider needs of the target populations?

'Network health': a P2P's ability to engage its members, sustain their engagement, and adapt as needed. May involve issues of trust building and management effectiveness.

Example question(s)

Has the P2P secured the necessary resources (capacities, money, and infrastructure) to become self-sustained? What are the network's governance rules and are they effective? Do decision-making processes encourage members to contribute and collaborate? How are the network's internal systems and structures adapting over time? Do all members share a common purpose for the network? Are all members working together to achieve shared goals, including goals that emerge over time? Are members achieving more together than they could alone? Has a sense of trust developed amongst the network participants?

'Network connectivity': the extent to which the members' ties to each other are resulting in efficient and effective "pathways" for shared learning and action.

Example question(s)

Has the P2P assembled members with the capacities needed to meet network goals (experience, skills, connections, resources)? Who is connected to whom? Who is not connected but should be? Is membership adjusted to meet changing network needs? What are the number, quality, and configuration of network ties? How dependent is the network on a small number of individuals? Is the network structure adjusted to meet changing network needs and priorities?

'Added value': changes that can reasonably be argued to be due to the P2P operation, rather than any other factors.

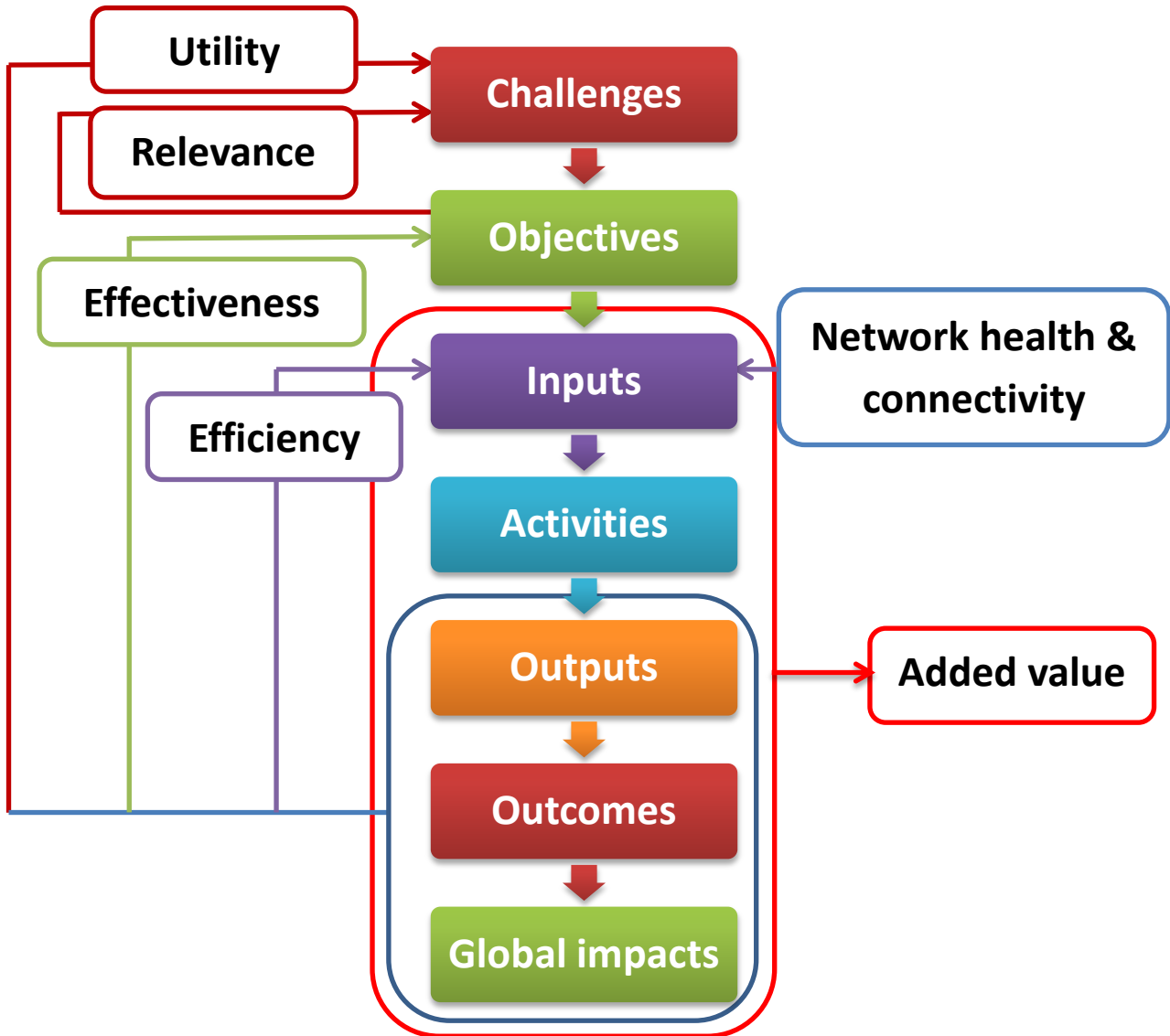
Example question(s)

What is the additional value resulting from the P2P, compared to what could be achieved by Member States alone at national and/or regional levels? To what extent do the problems/challenges addressed by the intervention require action at EU level? What would be the most likely consequences of stopping or withdrawing the existing policy intervention?

Source: amended from EC Better Regulation Toolbox, CIA4OPM (2011) and Network Impact and Centre for Evaluation Innovation (2014).

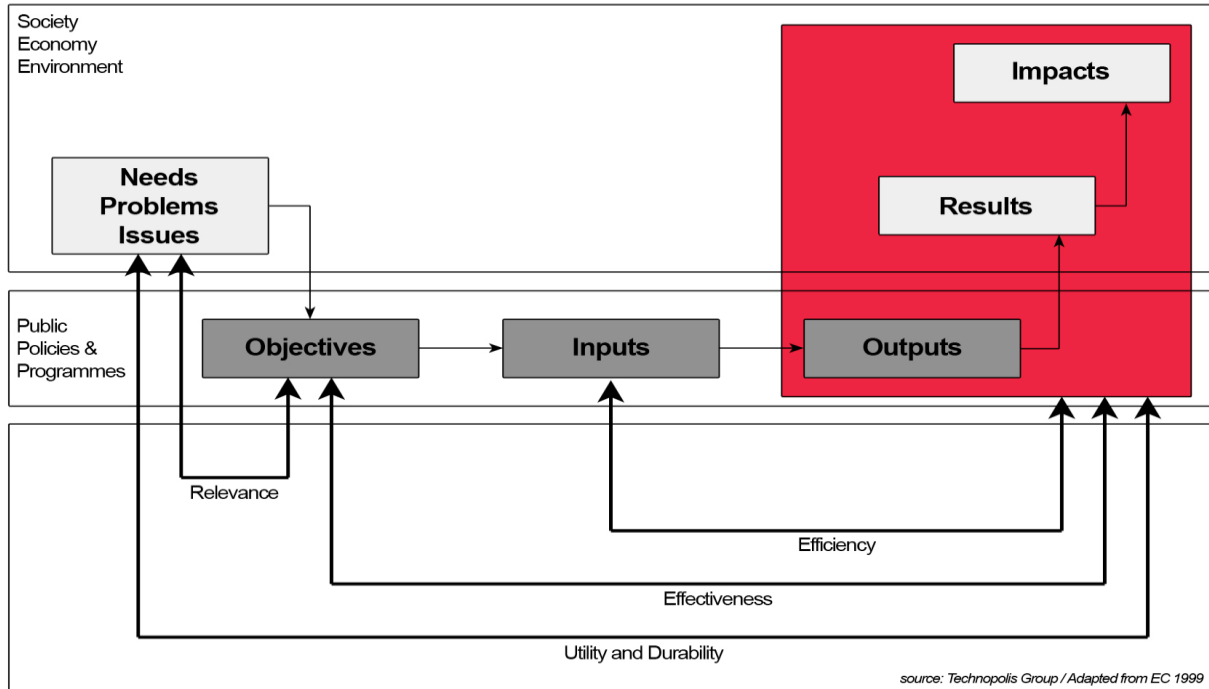
It is not unusual that evaluations focus only on some of the evaluation issues rather than trying to address all of them at the same time. In fact the most common issues that are usually addressed are effectiveness, efficiency, relevance and utility and added value. An illustration of how certain evaluation issues relate to the P2P elements is provided below.

Figure 7: Basic elements of an Impact Assessment Framework including some evaluation issues



A real case example comes from JPND as illustrated below.

Figure 8: JPND Framework for evaluation – linking evaluation issues to the JPND components



Source: Technopolis 2012, JPND Evaluation Framework

Setting the appropriate time for an evaluation

Evaluations can be carried out at different times, i.e. before, during or after a policy intervention has taken place. In each of these cases they serve different purposes.

Definitions

Ex ante evaluations are conducted **before** the intervention is implemented. The ex ante evaluation brings together evidence and arguments concerning the likely consequences of the intervention’s activities. This evaluation may challenge the clarity with which the objectives have been set out, as well as the plausibility of the objectives themselves. In a sense, it tests the assumptions underlying the design of the intervention and its anticipated results and outcomes against the rationale for its creation. Ex ante evaluation provides a valuable input to later – real-time and ex post – evaluations. It should create a knowledge base (a “virtual benchmark”) against which these evaluations can be undertaken.

‘Interim’ or **‘intermediate’** evaluations are conducted at some point during the implementation of a programme/intervention, at a point where some early results should be apparent. In general, early interim evaluations can be used to provide management feedback on the uptake and administration of the intervention, whilst later interim evaluations may focus more on the initial outcomes from the various activities it supports. This type of evaluation can inform the implementation of later stages of the programme/intervention, the selection of activities and sub-activities, and so on. It can play an important political role in providing legitimacy for ongoing long-term funding. It can provide early

warnings and timely information that can alert and inform programme managers of the need for action, for example where participants are experiencing difficulties, indicate where future underperformance might be expected, illuminate unexpected consequences of activities, identify where there are shortcomings in communication or challenges to morale, and so on.

Interim evaluation should not be confused with **monitoring** which is closer to a management process of checking that funds are being spent, activities accomplished according to time and that information on certain immediate outcomes or outputs is being captured. Monitoring, however, forms an important element of ex post evaluation as certain data and information that are important for ex-post evaluation can and should be collected during the monitoring process. Indeed, routine monitoring and data collection may save considerable resources during an ex post evaluation as they avoid the need to try to recapture data that may be partial or lost.

Ex Post evaluations are undertaken after a programme/intervention is **completed, or at least near enough to the end** that most activities can be examined and the more immediate outcomes assessed. Ex post evaluation allows for the intentions of a programme / intervention to be confronted with the realities that the programme activities have encountered and for questions about the effectiveness, efficiency, etc. to be addressed.

Source: Miles and Cunningham, 2006.

Naturally not all evaluation types are appropriate for the examination of all evaluation: it is often more appropriate to focus on certain issues at one point in the policy cycle rather than another. The following figure provides an indication of those issues which are best dealt with at each evaluation timing.

Figure 9: Evaluation issues or aspects and timing of evaluations

	Ex-ante evaluation/ impact assessment	Interim evaluation	Ex-post evaluation	Ex-post impact assessment
Relevance	X	X		
Coherence	X			
Economy	X	X	X	
Effectiveness	X	X	X	X
Efficiency	X	X	X	X
Sustainability			X	X
Utility	X		X	X
Consistency	X	X	X	X
Allocative/ distributional effects	X	X	X	X

Source: CIA4OPM, 2011, p. 37.

6. The importance of a monitoring system

In any impact assessment there is need for baseline data. This strongly suggests the need for a process of data collection at the application stage or very close to the project commencement stage. Much of this data can be collected as part of the normal set of questions and information sought on project participants.

The process of evaluation particularly that of *ex post* evaluation is greatly enhanced by the availability of robust and comprehensive information and data relating to project processes and outcomes. The availability of such data greatly facilitates the task of the evaluators and obviates much of the need to collect retrospective information, some of which may have been forgotten or lost as a consequence of participant mobility. In addition, the collection of such data *ex post* necessitates significant time and resources on the part of both the evaluators and the project participants. This underlines the advantages of implementing a process of monitoring which can be aligned with the normal reporting requirements. Through a 'light touch' monitoring process, a significant amount of information may be collated which will greatly assist the final evaluation process.

A suggested optimal period for monitoring/interim reporting for projects funded under ERA-NETs, Art. 185 or JPIs, is on an annual basis, 12 months from the project start. Since many of the project outcomes will not have materialised until the end of the project, a final reporting stage is useful to capture the complete set of project outcomes and any information relating to experiences directly related to the project itself. Typically, a final report form is used for this purpose.

Yet, certain outcomes and impacts from P2P networks or the funded projects are unlikely to be manifested until after the lifetime of the project itself, or after some time of the P2P existence. Thus, a further *ex post* reporting period can be desirable. A suitable trade-off is required to ensure that sufficient time has elapsed to allow the development of longer term impacts and that the participants' 'memory' is still sufficient to enable the collection of meaningful information. Thus, a period of between 6 months to one year following the end of a project is suggested, even though for some impacts it may be relevant to allow for a 3-5 year period running the risk, however, of limited 'project memory' and low ability of impact attribution.

A similar timetable for monitoring and assessment is useful to apply also at the network level. Baseline information may be collected at the moment the networks become operational (i.e. as soon as their administrative mechanisms are in place). Monitoring / interim information and data should, ideally, also be collected at regular intervals through the course of the networks, i.e. every 12 months, and again when the networks cease their operation. The end of the networks' operation however may not be the point in time when the EU support finishes. Some networks continue their operation even after the end of the EU-supported life cycle based on their own contributions. In these cases it would be good to gather monitoring and assessment information at the end point of the EU support period to have a reference point for the whole EU supported operation of the networks as well as a baseline reference for their "independent life".

Project level information collection

Application/project commencement phase

At this stage, it is very important that the appropriate project objectives are elaborated and understood. These are important not only for the first stage of the evaluation itself but also to underpin the derivation of appropriate indicators and performance measures which may be applied to the later stages of the monitoring and evaluation of ERA-NET, Art. 185 or JPI-funded projects. Thus, it is essential that proposals should provide certain types of information and contain clearly defined research (or related) objectives, targets and milestones. Contingent on the overall objectives of the project, the following types of information appear to be the most relevant to this phase:

- Information about intended networking;
- Information about academic/industrial relevance;
- Information about intended dissemination and follow-up research; and
- Information about availability of resources.

The specific instrument to be used at this stage would be a standardised proposal form, which can collect the required information in a structured manner. Typically data/information will be collected as a matter of routine on participant details. Further data may be collected relating to:

- Level of staff receiving support
- Sources/amounts of co-funding for the project
- Level of prior contact with other project participants
- Centrality of research project to core activities of organisations
- Etc.

Interim/monitoring

The generic types of information required during the on-going evaluation and assessment will need to be of a standardised format and structure and thus will have to reflect the generic objectives of ERA-NETs, Art. 185 or JPIs rather than the specificities of individual projects. At this point or phase of the project, data collection may be concentrated on collaboration and networking processes and on initial outputs, including, for example:

- Publications (articles, conference proceedings, books, book chapters, reports, grey literature, datasets, etc.)
- Conference/workshop attendances
- Project meetings
- Degree theses
- Products, process etc. (licensed/patented or otherwise)
- Student/staff exchanges
- Contributions to standards, public awareness, policy
- Further development of research networks, et c.

End of project evaluation

At this stage, the evaluation will be largely concerned with issues relating to networking / collaboration, outputs and dissemination. It is anticipated that the monitoring process should, by this stage, already have resulted in the collection of a substantial amount of information relating to the indicators defined earlier. However, it will be necessary to collect additional information through a further reporting exercise, structured around an end-of-project questionnaire. The combination of accumulated information and the targeted questionnaire should obviate the need, at the individual project level, for a larger scale final evaluation undertaken by a designated evaluation panel (including external experts) – although this may be desirable at the network level. This should reduce the resources required for overall evaluation tasks and provide an additional source of consistent, comparable data which could be used by any future “meta-evaluations” of the entire ERA-NET, Art. 185 or JPI mechanism.

In addition to the collection of the types of data and information already described above, additional qualitative and forward-looking data could also be collected. This data could include:

- Extent to which project objectives were achieved
- Planned or actual collaboration activities resulting from project (describe project, partners, etc.)
- Source(s) of supporting funding for above
- Reasons for lack of continuation
- Impact of project participation on own research activities in this theme
- Other effects
- Outputs (see Interim/monitoring, above)
 - Assessment of scientific/technical quality of outputs
 - Were outputs of national, European or wider significance?
 - Most significant outputs
- Assessment of overall participation costs/benefits
 - Most significant benefits
 - Alternative opportunities

Programme / instrument level information collection

Network commencement phase

Despite the differences among the network types (ERA-NETs, Art. 185 actions, JPIs, etc.) the information to be collected at the network level follows the same approach. The first type of information that has to be clearly articulated concerns the objectives of the network irrespective of its specific type (ERA-NET, Art 185 or JPI). The network objectives have to be clearly understood: these must be understood vis a vis the broader objectives of the specific instrument and the wider goals of the EU research and innovation policy. Clearly articulating the objectives of the specific network will help build up the objective hierarchy of the specific instrument (see section 2b above). This will also help build the intervention logic (see section 3 above) of the network to form a framework for its assessment.

Similar to the assessment needs at the project level it is also essential that proposals for networks provide certain types of information and contain clearly defined research (or related) objectives, targets and milestones. The following types of information appear to be the most relevant to this phase:

- Information about the networks (types of networks, funding sources, budgets, national contributions, funding modes, countries represented, network objectives, activities, thematic priorities and S&T fields addressed),
- Information about network partners (number, types, contact data, role of organisation in network, funding source of organisation for the participation in the network, etc.)
- joint activities (number, types of activities, types of research and research fields addressed, sources of funds, national budget, EU budget, other budget, funding mode)
- joint calls (number, types of research and research fields addressed, sources of funds, national budget, EU budget, other budget, funding mode)

This information can be collected in a standardised proposal form by the proposal coordinator.

Interim evaluation/monitoring

As already noted for the project level, the information that has to be gathered at the interim stage will have to reflect the generic objectives of ERA-NETs, Art. 185 or JPIs rather than the specificities of individual projects. Relevant information would address, for example:

- The network activities (those completed against activities scheduled) and their features (e.g. number of training events, participants, exchange visits, etc.)
- The network budget (absorbed against overall)
- Joint activities completed against planned activities and their specific features (number, types of activities, outputs of activities, etc.)
- Joint calls completed against planned and their specific features (such as number, number of proposals, accepted proposals and participants, etc.)
- Stages in networks development achieved
- Etc.

End of network evaluation

At this stage, the evaluation will be largely concerned with issues relating to the overall aims of the network and respective type (i.e. ERA-NET, or Art. 185 or JPI), e.g. coordination or alignment or integration of national / regional programmes, excellent science, industrial leadership and tackling societal challenges, etc.

The monitoring phase should have gathered a substantial amount of information that can be used as comparisons for the set of information that will be gathered ex-post. In addition the project level monitoring and assessment procedures will prove quite useful at this stage of network evaluation because of the wealth of information they will have produced. When aggregated, this could provide valuable insights about impacts at the network level. However, the type of information and data collected at the project level will have mainly addressed issues like scientific excellence or industrial

leadership. Issues such as programme effectiveness, efficiency or additionally or the degree of achievement of the broader network objectives like alignment and coordination of national / regional programmes are still to be addressed mainly at this level. Thus, information might be collected with respect to the following:

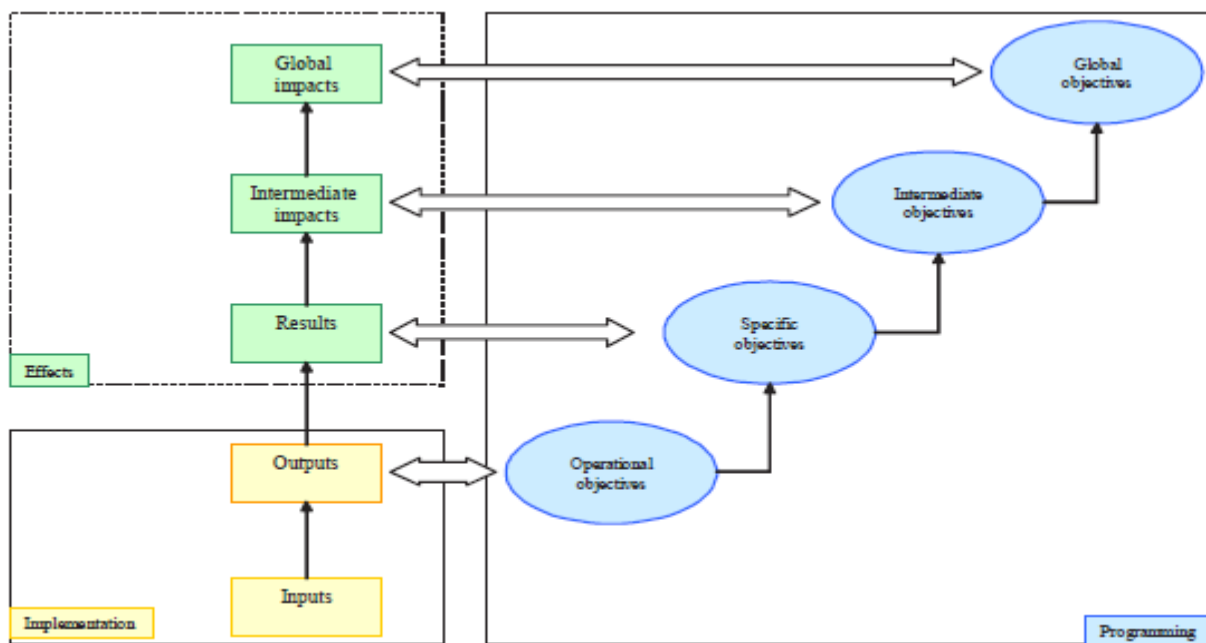
- Mutual learning:
 - Joint workshops, joint training, joint papers, exchange visits, identification of areas of common interest, SWOT analyses
 - Capacity building in national agencies
- Mutual opening up of existing programmes
 - Agreed rules, procedures, initiatives for mutual opening up of infrastructure / programmes
 - Agreed rules and procedures for joint access to research data, databases, etc.
- Scientific alignment
 - Programme clustering
 - Changes in research priorities of agencies
 - Alignment of national agendas
- Managerial alignment
 - Common programme monitoring and evaluation schemes
 - Harmonised rules and procedures for participation
 - Coordination of timing in funding and programme implementation
 - Multinational evaluation schemes
- Financial alignment
 - Changes in legislation to allow payments to foreign researchers
- Networking/collaboration of national agencies
- Joint design, implementation and monitoring of new programmes
- Mobilisation of national funds / resources for international research
- Etc.

The development of indicators to measure outputs, outcome and impacts at project and network levels is discussed in the next section.

7. Defining output, outcome and impact indicators

Measuring the effects of policy interventions is of major importance both for accountability as well as improvement purposes. Although there are several ways of capturing results and impacts, there tends to be a preference, especially among policy-makers, for the use of (generally quantitative) indicators, which often attain higher visibility in policy debates than do qualitative impact statements. Indicators need to apply to different outcome levels (short-term outputs, intermediate results, long-term impacts) in a similar way that the logic model establishes a hierarchy of linked objectives at different levels (operational, immediate, intermediate, and global). Since the intervention logic is highly likely to be subject to some evolution during the lifetime of the policy intervention due for instance to change of assumptions over time, it is important that indicators are also allowed to be revisited. (CIA4OPM, 2011)

Figure 10: Correspondence of effects (outputs, outcomes/results and impacts) with objectives



Source: European Commission, 2004.

It is also essential to understand that indicators are subject to a number of **limitations**. They cannot measure all aspects of the reality and are often just a proxy of an intended outcome, while indicators that are defined ex-ante can only capture intended impacts. The appropriateness of indicators is case and context dependent. Societal impacts appear especially difficult to measure but, in certain cases, the cost of developing a good information system may be as costly as the intervention itself. (CIA4OPM, 2011)

An interesting approach in developing indicators comes from JPND. In this case two types of indicators are proposed: Type A: these monitor the effect of JPND on (European) research programming, research policy and funding (the concept of joint programming) and Type B: these monitor the scientific and societal impact of JPND research on degenerative diseases. This separation follows the distinction of the

JPND logic frame to two perspectives, i.e. the joint programming perspective and the scientific and societal perspective (cf. Figure 4 above) and can be paralleled to the policy and project level. Following this separation, the two types are then set to include input, output, outcome and impact indicators. These indicators are described in detail along with suggestions on how and when to collect relevant data in sections 1.4.3.1 and 1.4.3.2 of the relevant Technopolis report available at http://neurodegenerationresearch.eu/uploads/media/Monitoring_and_Evaluation_Framework_1.8Mb_.pdf.

JPI MYBL has also followed a similar approach. The first external evaluation of JPI MYBL that was recently completed focused on mostly Type A indicators addressing process, input, output, and outcome indicators. The report is available at http://www.jp-demographic.eu/wp-content/uploads/2015/11/3a_D6.6_J-AGE_External-Evaluation_Final-Version.pdf.

Following its evaluation framework, JPI AMR suggests nine different types of indicators, i.e. three types of structure, process and outcome indicators for each of the three analysis dimensions of governing policy making, governing research performance, and responsiveness and innovation. The two levels of governing policy making and governing research performance can be paralleled to the network level while the third one might refer to the project level. The relevant report is available at <http://www.jpamr.eu/wp-content/uploads/2016/02/Framework-monitoring-evaluation-JPI-AMR.pdf>.

Other networks have also developed indicators and carried out their evaluations mainly at project level. Some examples along with the respective reports can be found at the Reference Library of ERA-LEARN 2020 available at <https://www.era-learn.eu/monitoring-and-assessment/reference-library>.

Building on all the above, some suggestions can be made on more 'generic' type of indicators to allow for diversity, which however can be adjusted to each particular case. The aim is to offer hints to think about what the proper indicators would be for a particular case from a variety of suggestions that may or may not have been considered in the absence of such suggestions. This is attempted in the two following tables (2 and 3) including a number of suggested output, outcome and impact indicators at both the network and project level. Naturally, these need to be adjusted to the objectives and activities of each specific P2P. There is also an attempt to separate outputs and impacts per different type of beneficiary (

Table 4).

Table 2: Examples of indicators at P2P network level

Activity	Sub-activity	Output Indicators / nature	Outcome Indicators / nature	Impact indicators / nature	Source of information	Timing
Mapping national/trans-national activities	Mapping workshops/ meetings	No of attendants (quant.) Quality of report/ deliverable (qual.) Programme clustering (qual.)	Identification of common areas of interest (qualitative)	Critical mass of research in certain areas (both quant. qual.)	Monitoring/ questionnaire	Interim/ ex-post
Foresight and common vision	Foresight exercise Vision building ws	No of attendants (quant.) Quality of report/deliverable (qualitative)	Identification of common areas of interest (qualitative)	Inform national and European policies (qualitative)	Monitoring/ questionnaire	Interim/ ex-post
Strategic Research Agenda / Implementation Plan	Interaction with AB, stakeholders Specific surveys	No of attendants (quant.) Quality of discussions (qual.) Quality / level of approval of SRA (qual.)	Identification of themes for calls (qual.) changes in research priorities of agencies (qual.) alignment of research strategies (qual.)	Specific strategies for certain areas (qual.) Influence national strategies/policies/ programmes (qual.) Changes in national budgets (quant.)	Monitoring/ questionnaire	Interim/ ex-post
Joint calls	Building a portal Call management Evaluation of prop.	User-friendliness of portal (quant. qual.) No of proposals submitted/ approved (quant.) Time to contract (quant.)	Promotion of research area at national levels (quant.) Change of national rules, timings (qual.) Multinational evaluation schemes (qual.)	Common rules, procedures, timing, and evaluation panels (qualitative) Changes in legislation to allow payments to foreign researchers (qual.)	Monitoring/ questionnaire	Interim/ ex-post
Research Infrastructures	Shared use of infra. Building new infra.	Hours of shared usage (quant.) Quality of joint use (qual.) New rules and procedures for new infra. (qual.)	Increased efficiency in use of infra (quant.) Increase use of infra (quant.)	Harmonised rules and timing in participation (quant. qual.) Harmonisation in research processes/ data (qual.)	Monitoring/ questionnaire	Interim/ ex-post
Stakeholder involvement	AB membership SRA involvement Proposal evaluation	Composition of governance structures (quant.) Share of industry /society in AB, other structures (quant.)	More relevant research to stakeholders' needs (qual.)	Increased stakeholder participation in national / European decision-making (quant.)	Monitoring/ questionnaire	Interim/ ex-post

Activity	Sub-activity	Output Indicators / nature	Outcome Indicators / nature	Impact indicators / nature	Source of information	Timing
Dissemination/ Uptake of research results	Website	No. of hits (quant.)	No of papers with new counterparts	Spin-out companies for commercialisation (quant. qual.)	Monitoring/ questionnaire	Interim/ ex-post
	Newsletter	No. of recipients (quant.)	New proposals in funding programmes (quant. qual.)	New collaborations (quant. qual.)		
	Academic papers	No. of conferences (quant.)				
	Brokerage events	No of highly-ranked papers No. of attendants (quant.)				
Widening participation	Conferences / ws	New partners (quant.)	Better success rates of widening countries in H2020 (quant.)	Increased research performance of widening countries (quant. qual.)	Monitoring/ questionnaire	Interim/ ex-post
	Signing of MoUs	Increased budgets (quant.)				
Internationalisation	Conferences / ws	New partners (quant.)	New international proposals (quant. qual.)	New collaborations (quant. qual.)	Monitoring/ questionnaire	Interim/ ex-post
Monitoring and evaluation (M&E)	Set up M&E system	Quality of M&E reports	Informed decision-making (qual.)	Improved engagement of P2P partners	Monitoring/ questionnaire	Interim/ ex-post
		Budget allocated (quant.)	Improved network operation (quant. qual.)	Engagement of new partners		
Capacity building in national agencies	Internal training, exchange visits	Number of events / schemes (quant.) Participants / attendants at events (quant.)	Increase of expertise in coordination / collaboration (qual.)	Improved rules and procedures at national level (qual.)	Monitoring/ questionnaire	Interim/ ex post
Network health and connectivity	Guidelines for collaboration	Formal/informal guidelines produced (qualitative)	Increased engagement in activities over time	Long-standing collaboration of partners	monitoring/ questionnaire	Interim/ ex post
	Governance structures	Memberships (quant. qual.)	Improved governance over time	Self-sustained P2P	monitoring/ questionnaire	Interim/ ex post
	Network sustainability	Resources availability (quant. qual.)	Future intentions re collaboration (qual.)	Changes in national budgets re international activities (quant.)	questionnaire	Ex post

Table 3: Examples of indicators at project level

Project activity	Output Indicators / nature	Outcome Indicators / nature	Impact indicators / nature	Source of information	Timing
Research collaboration	Publications (quant.) New staff, students, employees linked to project/theme (quant.) New methods, services, products (quant./qual) Co-authorships (quant.) New joint proposals/projects (quant.) New theories, practices, trans-disciplinary activities (qual.)	Changes to research programmes of organisations (qual.) Increased collaborations (quant.) Higher-research ranking (quant.) Increased reputation (qual.) Access to additional R&I funding (quant.)	New research trajectories / new areas of research (quant./qual.) Solutions to challenges (qual) international profile (quant./qual) Increased long-standing collabs (quant./qual)	Monitoring/ questionnaire	Interim Ex post
Research collaboration Academia – industry	Industry/HE co-publications (quant.) Prototypes of new methods/products/services (quant.qual) Patents, licenses, leasing, etc. (quant)	New methods/products/services (quant.qual) Spin-offs (quant./ qual) Market share figures (quant./qual) Commercial returns – turnover – employment (quant.) Reduced operating costs (quant.)	Solutions to challenges (qual.) Increased industry competitiveness (quant/qual) Improved business models (qual.)	monitoring/ questionnaire	Interim Ex post
Results dissemination Science-society	Raising awareness in society (quant./qual)	Change consumers behaviour (quant./qual)	More informed / concerned citizens (quant./qual)	Monitoring/ questionnaire	Interim Ex post
Research collaboration Dissemination to policy	Inputs to standards (qual.)	White papers, draft regulations (quant./qual) Changes in policies / regulations (quant/qual.)	Solutions to challenges (qual./quant) Improved policy-making (qual.) Improved service quality (qual) Reduced environmental impacts (quant.qual)	questionnaire	Ex post
Capacity building Training / knowledge transfer	Training schemes/activities (quant./qual.); Masters/PhD students (quant.); Conferences, workshops, seminars (quant./qual.)	Improved capacities at organisational level (quant./qual.) Changes to human resources Organisational changes (quant./qual)	Improved national capacity / performance in specific area (quant./qual) New practices for research organisation (qual)	Monitoring/ questionnaire	Interim Ex post

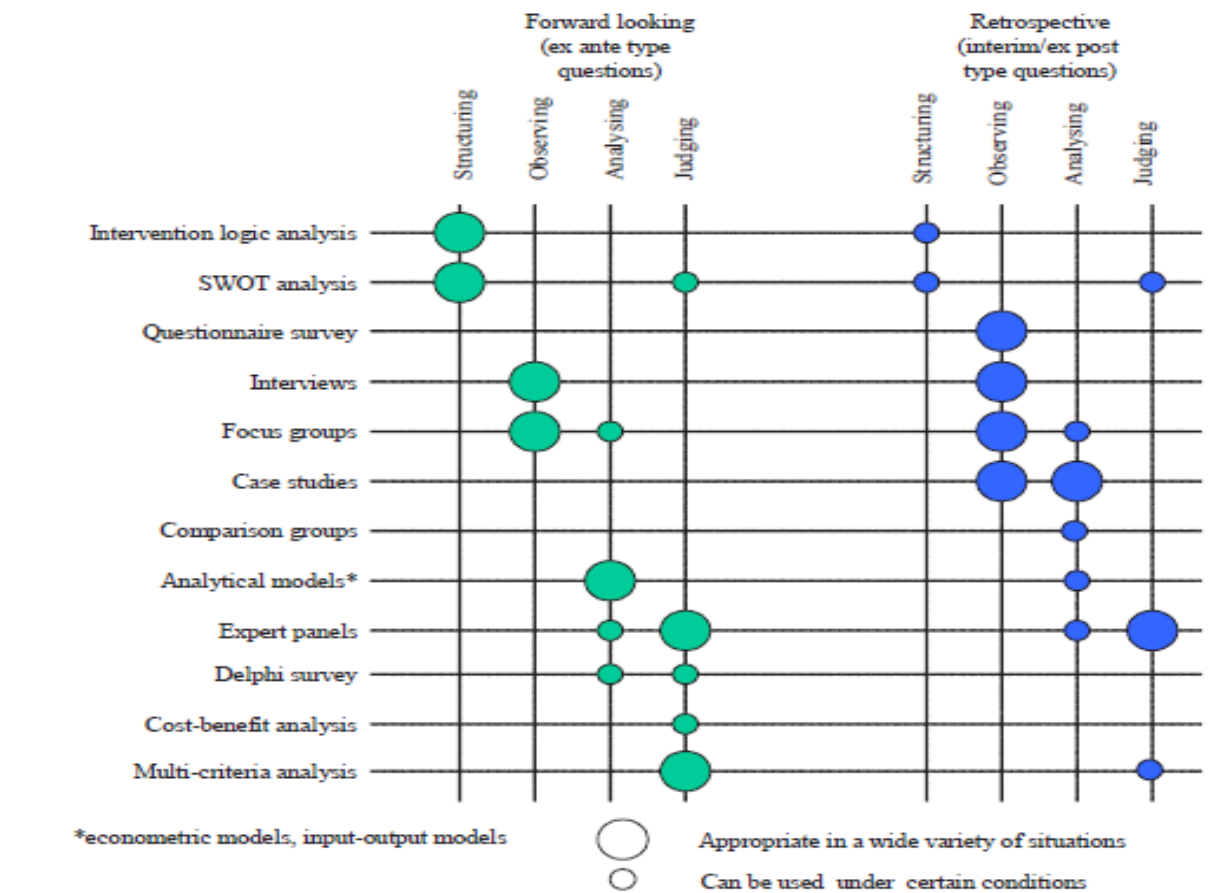
Table 4: Types of outcomes and impacts per different type of beneficiary

Type of Beneficiary	Outcomes	Intermediate Impacts	Global Impacts
Research organisation	new technology, new data/method, formal publications, patents	additional research income, commercial income, increased research capacity, spin-off businesses, enhanced reputation	new research trajectories, new solutions for socio-environmental challenges, economic spill-overs to industry
Industrial organisation	new product/service, new technical process, new organisational process, patent, improved capacities	increased turnover/profit, new jobs, protection of existing jobs, increased market share, geographic expansion	economic spill-overs to other businesses, new solutions for socio-economic challenges
Public service organisation	new methods/services, new organisational process	improved service quality, reduced cost of service delivery	improved health, safety, security and/or quality of life for citizens
Public administration	improved scientific evidence, new organisational process	improved governance, reduced administration costs, evidence-based policy making	improved economic, social and/or environmental impacts
Societal organisation	improved scientific evidence, improved services, improved capacities	increased influence	improved standards/regulations, improved quality of life
Environmental organisation	improved scientific evidence, improved services, improved capacities	Increased influence	improved standards/regulations, reduced environmental impacts

8. Data collection and analysis methods

There are a variety of data collection and analysis methods available to those conducting evaluations and impact assessment. However, their suitability for use is determined by a number of factors, such as the timing of the evaluation, the specific activities and actions being carried out and the nature of the measure being evaluated. The figure below corresponds various evaluation methodologies and approaches to the timing of an evaluation (ex-ante, interim, ex-post). For further details, interested readers may access a number of reports which provide guidance on which methodologies are most suitable for different types of interventions (such as science-industry cooperation, science parks, strategic research and technology measures, innovation funding or clusters).¹¹

Figure 11: Data collection / analysis methods and timing of evaluations



Source: adopted from MEANS guide in http://ec.europa.eu/agriculture/eval/guide/eval_activities_en.pdf

A short description of the various data collection and analysis methods for evaluation and impact assessment can be found in Annexes A.2 and A.3 of Technopolis Group and MIOIR (2012),

¹¹ See for example TECHNOPOLIS GROUP & MIOIR (2012): Evaluation of Innovation Activities. Guidance on methods and practices. Study funded by the European Commission, Directorate for Regional Policy.

http://ec.europa.eu/regional_policy/sources/docgener/evaluation/pdf/eval2007/innovation_activities/inno_activities_guidance_en.pdf as well as on pages 25-27 in the following report https://wellcome.ac.uk/sites/default/files/wtp052364_0.pdf.

Several networks have also developed and used questionnaires for surveys or interviews. Such examples are available at the ERA-LEARN 2020 reference library, <https://www.era-learn.eu/monitoring-and-assessment/reference-library>. Further, there is a variety of reports that describe how case studies to identify impacts were carried out and what the results were. For example see <http://www.esrc.ac.uk/files/research/research-and-impact-evaluation/research-impact-on-practice/>,

After analysing more than 200 evaluation reports, the INNO-APPRAISAL study¹² found some interesting links between certain data collection / analysis methods with the evaluation issues addressed:

- A first group of evaluations is concerned with policy and strategy development issues and looks at the evaluation issues of consistency and coherence. Those evaluations, use more often context, document and network analysis as well as before/after group comparison. Consequently, they are fairly strongly correlated with document search, focus groups and interviews. Policy development evaluations are also linked to cost/benefit analysis, indicating that the strategic decisions need some quantitative backing.
- Analysis of effectiveness and output, outcome and impacts very generally display a significant correlation with case studies, input/output and descriptive statistics. The data collection methods correlating with this cluster of topics are existing databases, monitoring data, interviews and participant surveys. The overall effectiveness thus relies on a mix of existing material and rather simple methods to be applied as a standard approach. In addition, general assessments as for outputs and impacts as well as the assessment of the quality of outputs also rely more on peer review.
- For the issue of additionality/added value, the methods and data collection approaches are slightly different. These are related with econometric and network analysis as well as counterfactual approaches. Moreover, for input and output additionality input/output analysis, before/after group comparison, control group and cost/benefit approach are significant. The data collection methods used for additionality topics are mainly surveys, monitoring data, interviews and document search.
- A further group of evaluations deals with efficiency issues. Both for programme and project efficiency, case studies and context analysis are important, linked with document search and focus groups, workshops, as it is essential to understand the management structures, processes and practices. Efficiency at the project level, quite logically, is also linked with more sophisticated methods (such as input/output analysis, cost/benefit approaches, network analysis and econometric analysis) that appear to draw on participant survey data.

An important aspect to consider is also the following: it is often the case that the same national agency/ministry takes part in a significant number of P2Ps. Retrieving data and information for each P2P separately using different templates and methods may lead to so-called 'evaluation fatigue', diminishing response rates. Thus, there is a clear benefit if a centralised system for gathering certain data across all P2Ps can be established.

¹² http://www.isi.fraunhofer.de/isi-wAssets/docs/p/de/publikationen/INNO-Appraisal_Final_Report_100228.pdf.

9. Communicating evaluation / impact assessment results

In a study conducted for the European Commission¹³, Cunningham and Gok (2010)¹⁴ identify the main issues that affect the degree to which evaluations provide usefulness and utility to policy makers. These primarily concern the information sought by the evaluation, namely:

- Information on the effectiveness of design
- Information on the effectiveness of management
- Information on the effectiveness of implementation
- Information on the effectiveness of the evaluation itself
- Information on the achievement of objectives
- Information on the broader impacts of the instrument

However, this set of issues is particularly relevant to evaluations of policy interventions that have, typically, been commissioned by those directly concerned with sponsoring or implementing the policy and thus have a narrow set of concerned stakeholders.

In the case of P2Ps, the set of stakeholders will be much broader and, at a minimum will include:

- Participating national government agencies
- National government sponsors
- The European Commission
- Participating research institutions
- Participating HEIs
- Participating scientists and researchers
- businesses
- Observer bodies
- Other interested parties
- The broader public
- Etc....

Each stakeholder group will have specific interests and needs concerning the outcome of any evaluation. Thus, funding agencies may wish to learn about the efficiency of the policy intervention to ensure that tax payers' money is being used optimally, whilst higher level policy makers will be interested in the networking effects and the extent to which other countries' agencies have become involved. Researchers will be interested in particular contributions to their research field, while the European Commission will wish to assess the overall contribution of the policy intervention towards its own set of objectives.

This varied set of stakeholder concerns will probably necessitate some tailoring of the evaluation outputs to the different needs of the evaluation audiences. However, the objectives hierarchy defined at a preliminary

¹³ INNO-Appraisal: Understanding Evaluation of Innovation Policy in Europe (2010)

¹⁴ Cunningham, P. and Gok, A. Chapter 4: Usefulness of Evaluations, in INNO-Appraisal: Understanding Evaluation of Innovation Policy in Europe (2010).

stage of the evaluation design can be used to distinguish between the various audience requirements since the policy objectives will reflect the information needs of the different stakeholders – provided it has been carefully constructed. Indeed, the extent to which the evaluation outcomes satisfy the informational requirements of the stakeholders can prove a useful test of the design of the objective hierarchy itself.

A further factor concerns the timeliness of communicating the results of the evaluation to the stakeholder groups. For instance, information on research outputs will be more useful to researchers and scientists if they are communicated rapidly since their information content can diminish rapidly with time. Hence, if information on publications is collected through routine monitoring procedures, the results of these could be communicated to participants on a regular basis, even prior to the conduct of the eventual evaluation. In contrast, longer-term effects, such as the sustainability of cost-sharing agreements and further joint research agreements will emerge at a later stage and will be of greater interest to policy makers from the relevant participating national agencies.

For these reasons, it can make sense to utilise a range of communication strategies for the different stakeholders concerned: not all stakeholders will be interested in receiving the full evaluation report (although all should be provided with access to it). Thus, it is suggested that specific parts of the report, relevant to the information requirements of different stakeholders, can be produced as stand-alone documents. Here again the objectives hierarchy can provide a useful guide to differentiating the elements of the evaluation, since certain groups of stakeholders will have an interest in the achievement of specific sets of objectives.

There are some good examples available of communicating impacts in a short, but comprehensive and attractive way:

- Environment impact report of EMRP, file:///C:/Users/effie/Downloads/EMRP_Call_2010_TP_Environment_Report_2016_08_23.pdf. EMRP also produces short, 2pp. impact cases based on specific projects.
- The JPIs produced a collective publication including a 1 p. description of the main achievements for each of the 10 existing JPIs. <http://www.jp-demographic.eu/about-us/jpi-brochure-2016/>
- The Academy of Finland produced a report on the impacts from their participation in ERA-NETs, http://www.aka.fi/globalassets/42julkaisut/eranet_report_final_yhd4.pdf.
- Other examples of reports can be found on <https://www.era-learn.eu/monitoring-and-assessment/reference-library>.

10. Some good practice tips

When preparing for an evaluation

- Be prospective: build in monitoring and evaluation in the project planning stage.
- There is no one 'right' way to do monitoring and evaluation. It is important to create space to think carefully about why we are doing things and find out what we really need to know. Take regular time to reflect on what your monitoring is telling you: doing this matters as much as any particular technical tool.
- Understand your stakeholder and audience requirements and expectations.
- Check any prior similar impact assessment to see what issues were addressed and what expectations were presented.

When organising an evaluation

- Carefully consider the budget and timeline of the evaluation - do not expect contractors to perform tasks that are too expensive for the given budget or too lengthy for the timeline.
- Remember the basic principles of monitoring and evaluation:
 - Involve stakeholders, particularly those who will use the results of the evaluation, from the start;
 - Agree with the donor or funder about outcomes;
 - Decide whether the evaluation will emphasise accountability to a donor (and evidence that plans were delivered) or emphasise learning (and reflect on what worked and what did not for future improvement).
 - Decide on objectives and/or outcomes and associated indicators and method.
 - Ensure monitoring and evaluation is properly resourced – financially and in terms of staff time.
 - Ensure it is practical, usable and proportionate
- Encourage consideration of the "end-user" perspective. End-users have practical experience of what has happened on the ground and may have a different perspective from policy makers, governments, NGOs etc.

When commissioning the evaluation

- When commissioning the evaluation to external contractors consider carefully the degree to which the choice of methods will be left to them or a particular approach will be specified in the Terms of Reference (ToR) or methods will be jointly specified. ToR should also draw the attention of potential contractors to a range of sources of information and ideas. Contractors should be asked to explain in their bid the advantages, limitations and risks involved in using the proposed tools and techniques.
- The ToR can ask contractors to set up a panel of independent academic experts to review the content and quality of their deliverables. The ToR should also describe how the results of the project will be used and specify that contracts can be discontinued whenever the quality of the deliverables is insufficient in light of the terms of reference and the set quality assessment criteria, and when the contractors have not taken the necessary steps to remedy the insufficiencies.
- The contractor may be requested to further elaborate the evaluation questions presented in the ToR and may suggest additional sub questions and should indicate success criteria, relevant indicators and the

sources for the indicators/methodology for gathering necessary evidence. The sub questions break down the overall questions into more manageable issues, and allow for a structured and logical response to the higher level questions.

When carrying out the evaluation

- Evaluation questions should be worded in a way that forces the evaluator to go beyond an answer based on simple description. Questions that start with How, Why, To what extent are more likely to ensure that the answer provided looks at what the links were between the changes observed and the policy intervention(s). Questions that start with verbs such as "Do" or "Are" or "Should" need to be avoided as they tend to provoke yes/no answers.
- Try not to have too many evaluation questions. Sometimes it is necessary to have very specific questions, other times it is better to have a more generic set and see where the data/analysis leads. There is always a trade-off between the number of questions that are set and the depth of analysis that can be conducted.
- Choose appropriate methods and tailor them: there is no right or wrong.

Overall,

- Be flexible and iterative. Learning is part of the process.

11. Sources / suggestions for further reading:

A collection of evaluation studies can be found at http://www.fteval.at/en/evaluation_studies/

CIA4OPM. 2011. Optimizing the research and innovation policy mix: The practice and challenges of impact assessment in Europe. European Commission, Belgium, ISBN: 9789077735152.

Cunningham, P. and Gok, A. Chapter 4: Usefulness of Evaluations, in INNO-Appraisal: Understanding Evaluation of Innovation Policy in Europe (2010). http://www.fteval.at/upload/INNO-Appraisal_Final_Report.pdf

Cunningham, P., and Ramlogan, R., 2012. The Effect of Innovation Network Policies. Compendium of Evidence on the Effectiveness of Innovation Policy Intervention, MIOIR. NESTA, March 2012.

DG DEVCO, Evaluation of Joint Programming http://ec.europa.eu/smart-regulation/roadmaps/docs/2016_devco_004_evaluation_roadmap_jp_en.pdf

ERA-LEARN 2 Deliverable D4.3 Monitoring and Assessment Framework for P2P Activities, <https://www.era-learn.eu/publications/other-publications/era-learn-2-report-deliverable-d4-3-analysis-of-options-for-future-platforms-monitoring-and-assessment-framework-for-p2p-activities>

European Commission, 2004. Evaluating EU Activities. A practical Guide for the Commission Services, DG Budget http://ec.europa.eu/agriculture/eval/guide/eval_activities_en.pdf

European Commission. Better Regulation Toolbox, http://ec.europa.eu/smart-regulation/guidelines/docs/br_toolbox_en.pdf

INNO-Appraisal: Understanding Evaluation of Innovation Policy in Europe (2010) http://www.fteval.at/upload/INNO-Appraisal_Final_Report.pdf

JPIs to Co Work, 2013. Deliverable n° D3.3. June 2013.

Miles, I. and Cunningham, P. (2006): A Practical Guide to Evaluating Innovation Programmes. Brussels: ECSC-EC-EAEC.

Network Impact and Center for Evaluation Innovation, 2014, Part 1 of a Guide to Network Evaluation. Framing Paper: The State of Network Evaluation July 2014.

OECD glossary of terms in Glossary of Key Terms in Evaluation and Results Based Management, <http://www.oecd.org/dac/evaluation/dcdndep/43184177.pdf>.

Technopolis 2012, JPND Evaluation Framework, http://www.technopolis-group.com/wp-content/uploads/2015/04/1504_JPND-Monitoring-and-Evaluation-Framework1.pdf

TECHNOPOLIS GROUP & MIOIR (2012): Evaluation of Innovation Activities. Guidance on methods and practices. Study funded by the European Commission, Directorate for Regional Policy.

The Evalsed Guide,

http://ec.europa.eu/regional_policy/sources/docgener/evaluation/guide/guide_evalsed.pdf

The Magenta Book 2011

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/220542/magenta_book_combined.pdf

Weiss, C. (1997) "Theory-Based Evaluation: Past, Present and Future" in D. J. Rog and D. Fournier (eds.), 'Progress and Future Directions in Evaluation: Perspectives on Theory, Practice and Methods', New Directions for Evaluation, No. 76. p. 68.

Wellcome Trust, International Public Engagement Workshop Report. Engaging with Impact: How do we know if we have made a difference, https://wellcome.ac.uk/sites/default/files/wtp052364_0.pdf.