



**Evaluation  
in the  
Youth Field:**



**Theoretical  
Framework  
for Evaluating  
Implementation  
of the  
Youth Goals**



**Ondřej  
Bárta**

## Disclaimer

This project has been funded with support from the European Commission. This publication reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained therein.



## Acknowledgements

It is vital to use feedback when dealing with a complex topic like the one presented here, and colleagues of mine who contributed to all stages of this publication have my gratitude. Thank you very much Helderise Rendall Évora, Laura Gies, Martin Kitzberger, Eleonora Kleibel, Josip Miličević, Nino Preložnjak, Linda Stein, Marios Philippou and Christiana Xenofontos!

# Contents

2	Acknowledgements
4	Introduction
6	<b>1. Evaluation: Definition and Key Steps</b>
10	<b>2. Involving Stakeholders in Evaluation</b>
	2.1 Who Are the Key Stakeholders and How to Identify Them?
	2.2 Examples of Typical Stakeholder Categories
	2.3 Conflict of Interest and Power Relationships
14	<b>3. Differentiating Types of Evaluation</b>
	3.1 Evaluation by Purpose
	3.1.1 Formative (Constructive) Evaluation
	3.1.2 Summative (Conclusive) Evaluation
	3.2 Evaluation by Phase
	3.2.1 Process Evaluation
	3.2.2 Outcome Evaluation
	3.3 Evaluation by Relationship
	3.3.1 External and Internal evaluation
	3.3.2 Independent Evaluation
	3.3.3 Empowerment Evaluation
	3.4 Timepoint of Conducting the Evaluation
	3.4.1 Ex Ante Evaluation
	3.4.2 Interim Evaluation
	3.4.3 Ex Post Evaluation
	3.5 Other Types of Evaluation
	3.5.1 Ongoing and One-shot Evaluation
	3.5.2 Collaborative (Participatory) and Objective Evaluation
	3.5.3 Goal-based and Goal-free Evaluation
	3.6 How to Choose an Appropriate Evaluation Type?
21	<b>4. Outlining Evaluation Methodology</b>
	4.1 Social Research Designs
	4.1.1 Quantitative Approach
	4.1.1.1 Questionnaire
	4.1.1.2 Open Source Quantitative Data
	4.1.2 Qualitative Approach
	4.1.2.1 Interview
	4.1.2.2 Focus Group
	4.1.2.3 Observation
	4.1.3 Mixed Method Approach
	4.1.3.1 Case Study
	4.1.3.2 Action Research
	4.1.3.3 Experiment and Quasi-experiment
	4.2 Specific Evaluation Designs
	4.2.1 Needs Assessment
	4.2.2 Efficiency Assessment
	4.2.2.1 Cost-Effectiveness and Cost-Minimalization Analyses
	4.2.2.2 Cost-Benefit and Cost-Utility Analyses
38	<b>5. Debating Evaluation Ethics</b>
42	<b>6. Reporting on Evaluation</b>
44	Conclusions
45	Sources

# Introduction



Evidence-based policy in general is considered a holy grail in today's policymaking. It is desirable to base the policies on sound data, to have solid knowledge on the situation of the given phenomenon before introducing, re-introducing, or adjusting policies in the given domain. Evaluation is one of the tools which can contribute to the evidence-based policy in terms of monitoring and assessing current or past conditions and performances, employing a wide range of social science methods. This paves way for argument-based debates on policies for the future.

Youth sector is no exception, striving for evidence-based approach enabling improvements on all policy levels. European Union Youth Goals are a prime example of such youth policy. Based on voices of almost 50,000 young people from across Europe and using sound methodology, their views are mirrored in 11 Youth Goals<sup>1</sup>, presenting areas important to young people and in need of tackling. The Youth Goals are explicitly covered in the European Youth Strategy 2019-2027<sup>2</sup>, and as such provide basis for the EU youth policy for the years to come. The Youth Goals cover wide areas and as such are, intentionally, only the first step on the way: it is necessary to map these areas, to identify what has been achieved and what still needs to be done; to assess which particular actions are efficient and contribute towards achieving the Youth Goals, and which are not; in short, to evaluate which current youth policies are beneficial and which need to be adjusted.

This publication brings together current evaluation theories with an ambitious purpose of making the evaluation as such into a more widely understood and subsequently also a more frequently used tool in the youth sector. Wide audience is necessary in order to achieve this goal: practitioners, policymakers, youth researchers, as well as wide public. Practitioners may find evaluation useful when critically approaching their own everyday practice in order to both collect evidence of quality work and improving when necessary. Policymakers may get inspired and put evaluation mechanisms in place as part of the policy implementation phase, therefore enabling the evaluation to serve as continuous tool to provide evidence and help optimize as well as assess the implementation process as well as the policy itself. Youth researchers may appreciate another viewpoint on the research methods they are already familiar with and utilize them in a different way than usual next time they design a research project. And the wide public, particularly young people as beneficiaries of the youth policy measures, may strengthen their understanding and critical thinking on existent policies and the debates around them.

<sup>1</sup> YOUTHCONF 2019.

<sup>2</sup> Council of the European Union 2018.

Evaluation is an evidence-based policy tool, however, does not only need to be used at the level of policies. It can be used at the level of concrete mechanisms, at the level of particular projects and endeavours, and at the level of single organizations and institutions. Apparently, evaluation can be applied in a wide variety of contexts and aimed at assessing various phenomena. For the sake of clarity, however, policy evaluation is used throughout this publication when referring to the object of an evaluation. Despite this reference, the publication can be used in various contexts as well, simply replacing the word “policy” with a desired evaluation object in question.

This publication presents a holistic overview of the evaluation process. After defining evaluation as such, evaluation stakeholders are discussed to highlight important actors in an evaluation process and different evaluation types are described in order to show the broad variety of different approaches. Methodology section follows and provides a basic understanding of potentially useful social science methods to be utilized in evaluation endeavours. Subsequently, ethical considerations are debated, underlining the complexity, nuances and implications of an evaluation process. Dissemination of evaluation results is tackled separately, emphasizing the importance of sharing the outcomes with key stakeholders.

# 1. Evaluation: Definition and Key Steps



In order to understand the uses and varieties of evaluation, a clear definition of the evaluation itself is necessary. The basic definition of evaluation is a systematic and objective approach of determining the merit or the worth of an object.<sup>3</sup> In other words, the evaluation strives to find out value of a given phenomenon, such as a policy, a concrete policy measure, or a single project; and it is using a scientific approach based on systematic and objective procedures. Most of the authors add that evaluation has a purpose of improving the evaluated object, such as an operation, implementation or outcome of a project, program or a policy<sup>4</sup>. Some authors<sup>5</sup> also point out that evaluation processes are based on social research methodologies and professional standards. Rossi, Lipsey and Freeman<sup>6</sup> define evaluation research “as a social science activity at collecting, analysing, interpreting, and communicating information about the workings and effectiveness of social programs.” At the same time Patton<sup>7</sup> specifies that evaluation is involved in efforts to assess needs, formulate policies, pass laws, deliver programs, manage people and resources, develop communities, change organizational cultures, educate students, intervene in conflicts, and more. Other authors<sup>8</sup> add that the current view of evaluation emphasizes the relationship between evaluation and program or policy implementation in all of their phases, combining evaluation and implementation into an intertwined process<sup>9</sup>.

In conclusion, these are the basic elements of any evaluation:

- systematic approach;
- social science methods;
- value determination;
- purpose of improving the object.

Therefore, **this publication understands evaluation as a process of systematically, and using social science methods, determining a value of a policy (object, program, project, etc.) with a goal to improve it.** As mentioned above, “policy evaluation” will be used for the sake of simplicity throughout the text; however, as mentioned in the definition above, all information in this publication are applicable also to any programmes or projects as well as for one-time activities for, with and by young people.

<sup>3</sup> Joint Committee on Standards for Educational Evaluation in Frechtling et al. 2010: 3; Rossi, Lipsey, Freeman 2004: 2; Sida’s Evaluation Group 2018: 11; United Nations Evaluation Group in United Nations Development Programme 2019: 3.

<sup>4</sup> Patton 2002: 10; Rossi, Lipsey, Freeman 2004: 2; Weiss 1998: 4.

<sup>5</sup> Wholey, Hatry, Newcomer 2010: 5-6.

<sup>6</sup> Rossi, Lipsey, Freeman 2004: 2.

<sup>7</sup> Patton 2002:10

<sup>8</sup> Frechtling et al. 2010: 3-4.

<sup>9</sup> Ibid.

<sup>10</sup> C.f. Frechtling et al. 2010: 15-31.

As any process, evaluation consists of variety of steps, can have many forms and use various methods, as well as include numerous stakeholders. **The evaluation process comprises of five main steps<sup>10</sup>:**

- a) **ensuring in-depth understanding of the policy** which is to be evaluated;
- b) **defining guiding questions** for the evaluation process, in measurable terms;
- c) **choosing appropriate evaluation type and social science method(s)** to collect the data;
- d) **collecting and analysing desired data** with the aim of answering the evaluation questions;
- e) **reporting findings** of the evaluation to the key stakeholders.

The first two stages of any evaluation are of utmost importance, since they create a basis for the rest of the evaluation process, and in case the initial steps are not taken with care, the whole evaluation process may lead to distorted or invalid findings. At the same time, unfortunately, these are the most varying ones and depend highly on the policy to be evaluated, making it difficult to provide detailed checklist in order to ensure these first vital steps are well taken care of. Nevertheless, there are certain points to keep in mind when it comes to **(a) ensuring an in-depth understanding of the policy** which is to be evaluated. A conceptual model can be created by finding out information on the following key points<sup>11</sup>:

<sup>11</sup> Frechtling et al. 2010: 15-31; modified by author of this publication.

- **inputs**, (i.e. various funding sources that provide support to the policy);
- **activities** (i.e. services and actions defined as part of the policy);
- **outputs** (i.e. products of the abovementioned activities);
- **outcomes** (i.e. changes that occur as results of the abovementioned activities);
- **context** (i.e. specific features of the policy that may affect its implementation and ultimate generalizability).

It is obvious, that all parts of a given policy need to be mapped. As an example, employment policy focusing on young people in rural areas can be taken. Inputs are mapped in terms of all financial sources dedicated to the implementation of such policy, e.g. from the state, regional and local levels as well as from any private sources. Activities are named, e.g. specific trainings and requalification programmes, outreach programmes and counselling services, etc. Outputs are defined, e.g. number of new training programmes established, numbers of participants in the programmes, number of people using the counselling services, etc. Outcomes are measured, e.g. the employment rates in the particular rural area, etc. Last but not least, context is no less important, since it defines the overall links to the external environment.

For example a policy tackling unemployment of young people in rural areas can have some specific aspects in comparison to general employment-related policies, and these need to be taken into account; at the same time, this specificity needs to be assessed in terms of generalizability: is the policy viable outside of this particular context, i.e. will such youth policy work across all rural areas in the country? There are policies which may work miracles in certain rural environments, and at the same time be completely malfunctional in others. This needs to be taken into account when conducting the evaluation, and it needs to be part of the evaluation findings.

In practical terms, this conceptual model can be assembled either using simple diagrams and creating particular boxes under the given categories, or a mind-mapping software may be utilized to reach this goal.

<sup>12</sup> Frechtling et al. 2010: 15-31; modified by author of this publication.

Next step comprises of **(b) defining guiding questions for the evaluation process**, and since evidence is to be collected, this needs to be done in measurable terms. There are certain key points to take into account when diving into this task<sup>12</sup>:

- **identifying key stakeholders** and audiences;
- **formulating potential evaluation questions** of interest to the stakeholders and audiences;
- **defining evaluation outcomes in measurable terms**;
- **prioritizing and eliminating questions**.

Evaluation stakeholders is a complex topic and is treated separately in chapter 2. Nevertheless, stakeholders as well as potential audiences need to be taken into account, since different questions may be important to different actors. Evaluation serving only as a basis for internal reflection in a given ministry department is necessarily based on different aims than an evaluation for the purposes of informing public of the results of a given policy.

Formulating viable questions needs to be done in cooperation with the stakeholders, and it is beneficial to start from a wide range of questions which are then eliminated based on whether the answers can be found in the frame of the given evaluation at all.

As an example, however interesting it would be to find out if the local youth policy tackling non-formal learning in rural communities across Upper Austria also works in the Australian Outback, it is well possible that the resources dedicated to the evaluation efforts are not sufficient to determine that. Similarly, evaluating an impact of a local youth club on political participation across the whole country may simply not be doable, although the question itself is not irrelevant in nature.

Defining evaluation outcomes in measurable terms often helps in eliminating questions which are not answerable at all in the given evaluation framework.

As much as it is possible for Eurostat to conduct large-scale surveys on various topics, and frequently publish the results, it may not be possible to implement such survey for stakeholders involved in evaluation of a mental health youth policy in small settlements around Porto, and therefore comparisons of the results from Porto region to those of the Lisbon region simply may not be possible.

At the same time, all questions posed should be in line with ethical standards of the given country or working environment. Especially in the youth sector, where minors may be subjects to various policies, it is necessary that the ethical questions are asked and sufficiently answered, before further steps are taken.

Even though determining an impact of a newly established e-policy would be best determined via an application in smart phones of the young people, sending direct information on the usage of the e-policy tools for the purposes of evaluation, ethical risks of such an approach are apparent. Since ethical considerations are important throughout the whole evaluation process, chapter 5 provides more insight into this particular topic.



Viable guiding questions for the evaluation process should therefore be:

- **based on the interest** of evaluation stakeholders;
- **doable** in the framework of the given evaluation (e.g. in terms of funds, human resources, or mandate);
- **measurable** in the framework of the given evaluation (e.g. in terms of funds, human resources, or mandate);
- **ethically sound**.

Subsequently, **(c) choosing appropriate evaluation type and social science method(s)** to collect the data is necessary. Evaluation types are introduced in chapter 3 to show the wide variety of the evaluation processes, and hint on their potential uses in different evaluation frameworks; and it is this practical usage that constitutes the most important reason for the evaluation types to be listed and explained. Methodological approaches in line with the social research traditions are outlined, in terms of basic introduction, in chapter 4. The reason for not diving deep into this topic is simple: there are plenty of methodological books around<sup>13</sup>.

<sup>13</sup> University libraries are usually a good source of such books, and so are online open source publications.

Once the methodology of the evaluation is in place, **(d) collecting and analysing desired data** with the aim of answering the evaluation questions need to be conducted. There is no specific chapter dealing with these issues in this publication, since the information can be found, similarly to the methodology-related literature, rather easily in specialized books.

After the data has been collected, processed, and analysed, **(e) reporting findings** of the evaluation to the key stakeholders occurs. Chapter 6 is dedicated to this topic, outlining the potential ways of reporting and disseminating the evaluation findings.

# 2. Involving Stakeholders in Evaluation

Choosing the right type of evaluation and outlining its methodology cannot be done without knowing who the key stakeholders are and what interests are driving them. Identifying stakeholders and their interests is one of important steps in the planning stage of evaluation in order to design an evaluation which brings as useful information as possible and is also well-planned with regard to engaging stakeholders into the evaluation process. At the same time, knowledge of the stakeholders is important in the reporting phase, because it is much easier to report results when the audience is known.

<sup>14</sup> Frechtling et al. 2010: 24.

<sup>15</sup> Rossi, Lipsey, Freeman 2004: 50

<sup>16</sup> Wholey, Hatry, Newcomer 2010:

As some authors<sup>14</sup> point out, involving stakeholders from the initial phases of the evaluation process may increase the credibility of the evaluation as such and the likelihood that the evaluation outcomes are used once the process is concluded. To gain maximum benefit from the stakeholder involvement, involving stakeholders early, continuously and actively is recommended by experts.<sup>15</sup> Evaluators are encouraged to make a plan stating which stakeholders will be involved, in what phase of the evaluation, and in what capacity each stakeholder will be involved.<sup>16</sup>

All this being said, the concrete involvement of stakeholders in the evaluation is always a matter of informed decision by the evaluator. As debated in chapter 3, there are evaluation types which assume little to no direct stakeholder involvement in order to achieve its goals. All in all, openness and transparency in stakeholder involvement is desired in all cases, as debated in chapter 5 when exploring ethical considerations of evaluation endeavours. In this chapter, who the stakeholders are is described; how to identify key stakeholders is presented; typical stakeholders are listed; and power relationships are discussed.

## 2.1 Who Are the Key Stakeholders and How to Identify Them?

Stakeholders in evaluation are those individuals, groups, or organizations who have a vested interest (a stake) in the evaluation results, and are therefore interested in how well a policy works.<sup>17</sup> Usually, there is no homogenous group of stakeholders, as multiple groups with different interests and influences can often be identified.<sup>18</sup> These groups may include people with decision-making authority, funders and sponsors, administrators and staff, as well as clients (participants) or intended beneficiaries.<sup>19</sup> In other words, **stakeholders are all individuals, organizations, or policymakers, who are connected to the given policy.** In order to identify stakeholders, it is important to outline criteria which need to be fulfilled in order to label individuals or groups as stakeholders<sup>20</sup> and involve them in the evaluation process. **Stakeholders are such entities which:**

- **have content knowledge** of the evaluated policy (staff, expert consultant);
- **offer diverse perspectives** and experiences related to the policy (community leaders, target population, partner organization...);
- **are affected by the policy** (program participants);
- **are in positions of influence** to relevant groups as politicians (respected community leaders, advocates);
- **are proposers of evaluation** (funders, directors);
- **are responsible for decisions** about the evaluation and policy (organizational leadership, program director).

Brainstorming is the most frequently used method in identifying the stakeholders. It is advisable not to limit the brainstorming to evaluators, if possible, but to also include experts from the given policy area. In later steps, in case the policy terrain is a complicated one, evaluator can also involve the stakeholders who have already been identified in further brainstorming in order to finalize the list of stakeholders.<sup>21</sup>

Practically speaking, to evaluate a policy aimed at cross-border mobility of young volunteers between two countries, firstly the evaluator and a group of experts from both countries would put together an initial list of stakeholders, using their knowledge of the policy and the context. Subsequently, these initially identified stakeholders would be contacted and asked to identify further stakeholders connected to the given policy.

In doing so, the bullet points above can easily be utilized, pointing the stakeholders in the right direction and provide relevant information for stakeholder identification. The process stops when the evaluator gets repeated answers, and no new stakeholders are identified through subsequent interviews.

## 2.2 Examples of Typical Stakeholder Categories

To be more specific, examples of typical key stakeholder categories are listed in this subchapter. All of the below-mentioned categories of stakeholders, or just a few of them can be involved in the evaluation and an evaluator must be aware of their interests and concerns when including stakeholders into evaluation.<sup>22</sup>

<sup>17</sup> Chen 2015: 16; Rossi, Lipsey, Freeman 2004: 30.

<sup>18</sup> Chen 2015: 16.

<sup>19</sup> Rossi, Lipsey, Freeman 2004: 30.

<sup>20</sup> W. K. Kellogg Foundation 2017: 95-96.

<sup>21</sup> Cf. Bryson 2004a, Bryon 2004b in Wholey, Hatry, Newcomer 2010: 35-41.

<sup>22</sup> Rossi, Lipsey, Freeman 2004: 49.

- **Policymakers and decisionmakers** are responsible for deciding whether the policy is to be started, continued, stopped, expanded, restructured, or curtailed. Depending on the level of policy, these may involve governmental officials, local mayors and city boards, and other relevant individuals and structures.
- **Policy sponsors** are organizations that initiate and fund the policy. These stakeholders may overlap with policymakers and decisionmakers but can also be independent of them. In case of youth mobility related policies, for example, the policymakers and decisionmakers may belong to a national (governmental) level, but EU funds and programmes may be involved in the policy implementation, adding an international dimension.
- **Target participants** are persons, groups, households or other units that are recipients of the evaluated policy. Health related youth policy, as an example, may either be focusing on all young people in the given age range and locality, or may be tackling specific subgroups of young people, such as young people with certain health problems.
- **Policy managers** are responsible for overseeing and administering the policy. In the youth sector, this group depends on the structures in place. In certain situations, these will be personnel from an institution focusing on youth directly, in some cases, these may come from other areas. In the abovementioned example quoting health related youth policy, the policy managers may come from the public health sector.
- **Policy staff** are responsible for delivering the policy services. In the youth policy field, these are often NGOs or youth organizations, but the list may also go beyond these usual actors.
- **Policy competitors** are organizations or groups that compete with the policy for available resources. In case of youth policy focusing, e.g., on local housing, the policy competitors are all actors having stakes in city development, such as private companies and investors.
- **Contextual stakeholders** are organizations, groups, and individuals sharing the niche of a given policy, and therefore interested in the policy processes and achievements. In case of youth policy on, e.g., non-formal education, such contextual stakeholders can be universities which prepare youth workers for their professional engagement, or schools as institutions which occasionally also use the non-formal methodology in some of their activities.
- **Evaluation and research community** consist of evaluation professionals who read evaluation reports and pass judgment on their quality and credibility as well as of research professionals from the field(s) related to the evaluated policy.<sup>23</sup> Youth researchers are among the interested actors in the youth field, but professionals from other specific areas may get involved as well, given the cross-sectoral nature of some of the youth policies.

<sup>23</sup> Rossi, Lipsey, Freeman 2004: 48-49.

It is apparent from the example list above that the stakeholders may create a complex system of relationships, and these links may also be influencing the policy in question or the evaluation process itself. Evaluator needs to be aware of this complexity, as it brings in a potential for conflict of interest and a need to navigate a potentially comprehensive network of power relationships.

## 2.3 Conflict of Interest and Power Relationships

**Conflict of interest describes a situation in which multiple interests of one or several actors oppose each other.**

As an example, an individual in a role of a policy recipient who is at the same time designated as an evaluator of this policy can find themselves in conflict of interests. Some findings of the evaluation may suggest discontinuing of or adjusting the given policy as the best course of action, while some interests of the recipient may speak for the policy to stay in place and intact. In this situation, the conflict of the role of an evaluator and the role of a policy recipient is obvious. Nevertheless, much more covert and harder to find conflicts of interests may occur.

The conflict of interest is briefly presented here and options for dealing with power relationships are outlined.

Conflict of interest can range from an ethical issue (e.g. involuntarily involving one or several stakeholders in the evaluation process to a larger extent than others, based on personal relationships with these stakeholders), to behaviours which are in direct violation of laws (e.g. knowingly present evaluation results in a certain way in order to support certain financial interests). Typically, the conflict of interest can be either **financial** or **non-financial**, in the first case funding questions are concerned, while in the second a wide variety of other aspects can be involved: reputation, professional favours, personal relationships, etc. Any of the abovementioned conflicts of interest have potential to change an objective treatment to a preferential one, possibly affecting the evaluation processes and outcomes.

In any case of conflict of interest, it is important to identify these conflicts and inform about them from the beginning of the evaluation.<sup>24</sup> The basic rule is rather simple<sup>25</sup>: “*Conflict of interest should be dealt with openly and honestly, so that it does not compromise the evaluation processes and results*”. Once identified and communicated, the impacts of such conflicts of interests can be mitigated through implementation of various measures. Interviews with a stakeholder who is in a personal relationship with the evaluator can be conducted by an independent third party; financial conflicts may be resolved by full transparency of funding in the matters in question; and the whole evaluation can be reviewed by an independent expert at a certain stage, ensuring all sources are represented equally and without bias. These are only examples of possible measures to be taken, as this is such a complex topic that no universal checklist on such measures can be provided. In any case, it is advisable for the evaluators themselves to report their own identified conflicts of interests, together with measures taken to prevent them from affecting the evaluation process, in the evaluation report for the sake of complete transparency.<sup>26</sup>

<sup>24</sup> Cf. Weiss 1998: 111.

<sup>25</sup> The Joint Committee on Standards 1994 in Weiss 1998: 111.

<sup>26</sup> Weiss 1998: 111.

# 3. Differentiating Types of Evaluation

There are various types of evaluation which can be differentiated based on the following criteria:

- **purpose** the evaluation is to serve (aiming to conclude or to construct);
- **phase** of the policy the evaluation focuses on (stressing the process or the outcome);
- **relationship** of the evaluator and the policy (conducting internally or externally);
- **timepoint** of conducting the evaluation exercise itself (conducting evaluation before the policy implementation commences, during the implementation phase, or after the policy implementation is concluded);
- **other specificities** of the given evaluation approach (ongoing or one-shot; objective or participatory; goal-based or goal-free; problem or non-problem).

All of the aforementioned evaluation types are tackled in detail in the following subchapters. The aim of this exercise is not to complicate matters, but to offer an overview of viewpoints which are potentially useful to consider before starting to plan any evaluation process. At the same time, even though differentiated into subchapters for the sake of clarity, these types are not exclusive to each other, but can also be used to well complement one another, as will be shown in subchapter 3.6.

## 3.1 Evaluation by Purpose

There are two main types when dividing evaluation by its purpose – **formative (constructive) and summative (conclusive) evaluation**. An easy way to present the distinction between the formative and the summative evaluation is Snake's<sup>27</sup> famous illustration with soup tasting: *“When the cook tastes the soup, that’s formative evaluation; when the guest tastes it, that’s summative evaluation.”* Although this example cannot cover all evaluation activities<sup>28</sup>, it represents an easy way to understand the basic distinction between the formative and the summative approach.

<sup>27</sup> Snake in Chen 2015: 8.

<sup>28</sup> Chen 2015: 8-9.

### 3.1.1 Formative (Constructive) Evaluation

“Formative evaluation fosters improvement of ongoing activities”, states one of the authors<sup>29</sup>, stressing the main trait of such evaluation type. **Formative evaluation serves the purpose of improving the evaluated policy by forming it.**<sup>30</sup> Formative evaluations may relate to the sole necessity of existence of the given policy, to the design of the policy, its implementation process, its impact, or its efficiency.<sup>31</sup> Audience of the formative evaluation are typically stakeholders with an interest in optimizing the effectiveness of the policy.<sup>32</sup>

<sup>29</sup> Scriven in Chen 2015: p7.

<sup>30</sup> Patton 2002: 219; Rossi, Lipsey, Freeman 2004: 34; Wholey, Hatry, Newcomer 2010: 8.

<sup>31</sup> Rossi, Lipsey, Freeman 2004: 35-36.

<sup>32</sup> Ibid.

<sup>33</sup> Directorate-General for Education, Youth, Sport and Culture & ICF 2017.

<sup>34</sup> European Commission 2018a & 2018b.

An example of a formative evaluation is the Mid-term Evaluation of the Erasmus+ Programme (2014-2020), tendered by the Directorate-General for Education, Youth, Sport and Culture (European Commission), and conducted by a subcontractor, providing the European Commission with a final report<sup>33</sup> including recommendations, in order to summarize the achievements of the Erasmus+ Programme during the period preceding the mid-term evaluation process as well as to highlight potential for further positive developments in the upcoming periods. This report then served as a basis for further discussions, leading to a communication<sup>34</sup> from the European Commission to other responsible bodies, which included a summary of the main findings as well as suggestions for further actions in developing the Erasmus+ Programme.

<sup>35</sup> Directorate-General for Education, Youth, Sport and Culture 2017.

Another example of a formative evaluation is the Impact Assessment of the European Voluntary Service<sup>35</sup>, conducted by the Directorate-General for Education, Youth, Sport and Culture of the European Commission. Despite its name, the clear aim of the aforementioned evaluation was to assess selected aspects of the European Voluntary Service (EVS) in order to put forward recommendations to further develop the EVS itself.

### 3.1.2 Summative (Conclusive) Evaluation

Summative evaluation is any evaluation endeavour without formative elements<sup>36</sup>. It measures program outcomes and impacts during ongoing operations or after program completion<sup>37</sup> and **the purpose of summative evaluation is to provide an overall judgment about the effectiveness of an evaluated object (program) – to sum it up.**<sup>38</sup> Therefore this type of evaluation can usually serve the decision makers<sup>39</sup> at the end of a policy timeline when deciding if the evaluated policy should continue or not and in what form<sup>40</sup>.

<sup>36</sup> Scriven in Chen 2015: 8.

<sup>37</sup> Wholey, Hatry, Newcomer 2010: 8.

<sup>38</sup> Patton 2002: 218-219.

<sup>39</sup> Scriven in Chen 2015: 8.

<sup>40</sup> Patton 2002: 218-219.

<sup>41</sup> Chiwara & Ali 2018.

Examples of summative evaluation can be seen in the United Nations Development Programme. Each of the funded projects submits a summative evaluation at the end of the project cycle. As a concrete example, Joint Programme on Youth Employment evaluation<sup>41</sup> can be quoted. The publication summarizes and evaluates the main goals of the given project, and even though the recommendations are stated as well, these are aimed at external stakeholders, such as the government, in order to utilize the achievements of the project itself. The recommendations do not aim at the evaluated project itself; they provide information for a potential follow-up or a similar project conducted in the future. As described above, the aim of such a document is to provide basis for a conclusion and assessment of the given project, not for its continuation or further development.

Both the formative and summative evaluations can focus on process as well as on outcomes, as explored below.

## 3.2 Evaluation by Phase

Whereas formative and summative distinction relate to the objectives of the evaluation (forming or summarizing the evaluated policy), process and outcome evaluation types distinguish processes introduced in different phases of the policy implementation lifecycle.<sup>42</sup>

<sup>42</sup> Weiss 1998: 32.

### 3.2.1 Process Evaluation

Process evaluation is conducted continuously during the policy implementation and typically is focused on aspects such as participant enrolment, activities offered, actions taken, staff practices, or client actions.<sup>43</sup> **The aim of process evaluation is to assess the procedural aspects of a policy**, and can easily be included in both the formative and summative evaluations described above.

<sup>43</sup> Weiss 1998: 32.

An interesting example of a process evaluation is provided by Backett-Milburn and Wilson<sup>44</sup> in their insights into the concrete peer learning processes within health promotion work in the UK.

<sup>44</sup> Backett-Milburn & Wilson 2000.

### 3.2.2 Outcome Evaluation

After the policy implementation process is concluded, an outcome evaluation takes place, concerned with results and impacts of the policy, with the main emphasis put on what outcomes the policy shows in its target group upon its completion.<sup>45</sup> **The aim of the outcome evaluation is to look into the impacts and outcomes of a given policy** when the policy is concluded and the impacts and outcomes can be summarized and assessed.

<sup>45</sup> Weiss 1998: 32.

A rather famous Erasmus Impact Study published by the European Commission<sup>46</sup> and summarizing the impacts of the EU student exchange scheme. The report out of which the no-less famous “one million babies” finding was derived, is an example of the outcome evaluation: the aim was to determine any and all potential outcomes of the given policy, in this case, of the student exchange scheme.

<sup>46</sup> European Commission 2014.

## 3.3 Evaluation by Relationship

Evaluation can also be differentiated by taking into account the relationship between the evaluator and the policy itself. The following types can be identified and are described in the following subchapters: external (independent) and internal evaluation, independent evaluation, empowerment evaluation, and collaborative (participatory) evaluation.



### 3.3.1 External and Internal evaluation

The main difference between these two types lies in independence of the evaluator.<sup>47</sup> As one of the authors<sup>48</sup> explains: **“Internal evaluators are employed by an organization and are responsible for evaluating the organization’s own programs. External evaluators are not employees of the organization but are experts hired from outside to evaluate the program.”**

<sup>47</sup> Chen 2015: 14-15

<sup>48</sup> Ibid.

<sup>49</sup> Chen 2015: 15; modified by author of this publication.

There are aspects<sup>49</sup> which can be important in choosing internal or external evaluation, depending on conditions under which the evaluation is to take place as well as the general aim of the evaluation itself:

**Table 1:** Important aspects of internal and external evaluation.

Internal Evaluation	External Evaluation
Costs of the evaluation process are of great concern.	Cost of hiring an external evaluator is manageable.
Internal capacity/resources are available.	External evaluator can be hired.
The evaluator’s familiarity with the program is important.	Independence and objectivity are essential.
The policy is rather straightforward.	The policy is large or complicated.
Evaluation is for the purpose of monitoring or is constructive in nature.	The evaluation will focus on conclusive assessment.
Insider understanding is important.	Comprehensive assessment or fresh insight is needed.

Source: Chen (2015: 15; modified by author of this publication).

Internal evaluation can be, and often is, part of implementation structures, such as schools, youth centres, or NGOs.

An example of such permanent internal evaluation body is an Internal Evaluation Board of the Charles University, Prague, Czech Republic which aims to “support and develop quality assurance and internal evaluation of the educational and scientific, research, development, innovative, artistic, or other creative activities, and related activities of the University.”<sup>50</sup> As is apparent from the quotation, the aim is to monitor and further develop internal processes, and as such, to have internal experts and a continuous operation can be very advantageous. The ongoing internal evaluation is supported also by some governments, and useful tips for conducting such evaluation can be found in a publication by Education Review Office of the New Zealand Government.<sup>51</sup>

<sup>50</sup> Charles University 2018.

<sup>51</sup> Education Review Office 2016.

External evaluation, on the other hand, often comes to play in case unique conditions arise and the internal evaluation is not possible, or a conflict of interest might threaten its reliability and validity.

External evaluation of the Youth Peace Ambassadors Project<sup>52</sup> can be considered a good example of such endeavour: the project itself was implemented by the Council of Europe, and an external view was considered an imperative in order to assess the project, come up with new ideas and recommendations.

<sup>52</sup> Council of Europe 2014.

### 3.3.2 Independent Evaluation

In this type of evaluation, the evaluator is the only entity responsible for developing the evaluation plan, conducting the evaluation, and disseminating the results.<sup>53</sup> This is very important, as it is the main difference which distinguishes this type of evaluation from any other. **The evaluator can direct the evaluation absolutely autonomously and such evaluation does not need to be commissioned by the stakeholders of the policy itself** – it can be conducted by an independent research body or by an NGO interested in effectiveness of a new governmental youth outreach program.

<sup>53</sup> Rossi, Lipsey, Freeman 2004: 51.

<sup>54</sup> Independent Evaluations Office 2019.

<sup>55</sup> Ibid.

An example of such evaluations taking place can be found in the Independent Evaluations Office of the United Nations<sup>54</sup>. This body acknowledges that “*evaluation is critical in helping countries achieve the simultaneous eradication of poverty and significant reduction of inequalities and exclusion. By generating objective evidence, evaluation helps UNDP achieve greater accountability and facilitates improved learning from past experience.*”<sup>55</sup>

Independent evaluation is a tool which can not only help to develop policies based on independent results, but also help to balance the power relationship between the policymakers and the recipients of the policy by conducting a critical and objective research into the parameters of the policies in place and gathering evidence which can be subsequently used in an ongoing public debate.

### 3.3.3 Empowerment Evaluation

**Empowerment evaluation is designed and conducted by community members** while the evaluator’s role is to teach them basic evaluative techniques, coach them, and provide assistance when needed.<sup>56</sup> Rossi et al.<sup>57</sup> highlight that “*empowerment evaluation most appropriately involves those stakeholders who otherwise have little power in the context of the program, usually the program recipients or intended beneficiaries*”. Thus, the objective of this type of evaluation is to legitimize experiential knowledge of community members and empower them by democratizing research inquiry and at the same time improving the relevance of evaluation data for communities.<sup>58</sup>

<sup>56</sup> Weiss 1998: 99.

<sup>57</sup> Rossi, Lipsey, Freeman 2004: 51.

<sup>58</sup> Fawcett et al. 1996 in Weiss 1998: 99.

In this case, empowerment and direct experience of the recipients of the policy are the main aims and the evaluator acts as a guide along the way, making sure the methodological and processual requirements are observed and the evaluation outcomes are reliable and valid. At the same time, throughout the evaluation process itself, public participation is encouraged and enhanced.

An in-depth review of the empowerment evaluation was put together by Miller and Campbell<sup>59</sup>, including detailed description of various aspects of the empowerment evaluation itself.

<sup>59</sup> Miller & Campbell 2006.

Such evaluation is most suitable for local community-centred policies which have rather limited scope and the empowerment evaluation is viable.

Examples of such evaluations are brought by Newell and Graham<sup>60</sup>, quoting their experience with such methodology used in the Australian context.

<sup>60</sup> Newell & Graham 2012.

## 3.4 Timepoint of Conducting the Evaluation

It needs to be noted that there are terms related to the timing of evaluation efforts: ex ante evaluation; interim evaluation; and ex post evaluation. Each of these types is specific in itself as well as useful in combination with some of the evaluation types mentioned in other subchapters.

### 3.4.1 Ex Ante Evaluation

<sup>61</sup> Cf. Kloosterman, Giebel, Senyuva 2007: 30.

**Ex ante** (also called initial evaluation<sup>61</sup>) **stands for an evaluation conducted ahead of the policy implementation itself**, usually including SWOT analysis, needs analysis and other methods determining potential benefits as well as challenges of the policy in question. Aims of this particular evaluation are to prepare such implementation and monitoring mechanisms as to achieve the most efficient policy with high quality outcomes.

Ex ante evaluation has been conducted in 2004 as part of the preparations for the Youth in Action programme to be established<sup>62</sup>, and based on inputs from various stakeholders and experts brought together insights into potential implementation strategies of the upcoming youth mobility programme.

<sup>62</sup> Commission of the European Communities 2004.

### 3.4.2 Interim Evaluation

Interim, or mid-term, evaluation is an evaluative process which is conducted at such a point of policy implementation which allows outcomes of the implementation process to be seen, although it is too early for an overall, final evaluation. **Interim evaluation is usually conducted half-way through the policy implementation process** but can also be conducted at other timepoints, provided that the policy implementation has not finished, and outcomes are already detectable.

An example of this evaluation type is the Youth in Action Interim Evaluation, and exercise conducted in order to explore Youth in Action mobility programme for young people in 2011<sup>63</sup>. The aim of the evaluation was to provide both an insight into the then current mobility programme and its functioning, and at the same time to bring in recommendations for a subsequent mobility programme for young people<sup>64</sup>. This example well illustrates that elements of other evaluation types can be combined in this evaluation type (as well as in many others): formative and outcome evaluation meeting in one exercise.

<sup>63</sup> McCoshan et al. 2011.

<sup>64</sup> McCoshan et al. 2011: 11.

### 3.4.3 Ex Post Evaluation

<sup>65</sup> Cf. Kloosterman, Giebel, Senyuva 2007: 31.

**Ex post**, also called final<sup>65</sup>, **evaluation designates an evaluation process conducted at the end of the policy implementation** and as such takes into account all aspects of the given policy which took (or should have taken) place during the whole policy implementation period.

<sup>66</sup> Commission of the European Communities 2008.

Final evaluation of the “Youth” community programme<sup>66</sup> is an example of such endeavour, providing an overview of all elements of the programme, including all financial and statistical indicators, and drawing conclusions on the overall youth mobility programme outcomes.

## 3.5 Other Types of Evaluation

Evaluation types based on further specificities are listed below. Since this publication aims at providing initial information on evaluation processes, the most frequently mentioned types were chosen to be presented here. It is noteworthy, however, that there are more types of evaluation than mentioned in this publication, and some authors even introduce their own typologies.<sup>67</sup>

<sup>67</sup> Cf. Chen 2015.

### 3.5.1 Ongoing and One-shot Evaluation

<sup>68</sup> Patton 2002: 9.

Timing is the determining specification in these two evaluation types. As Patton<sup>68</sup> describes: “*The timing of the evaluation can range across continuum from a one-shot study of a specific aspect of implementation or one set of outcomes to an ongoing assessment system.*” Apparently, this specificity can occur in any of the aforementioned evaluation types, either building up on continuous data inputs and assessments, or working with a single data collection. None of these approaches is generally more suitable than the other, it always depends on the overall context of the evaluation as well as on the opportunities at hand: sometimes a one-time survey at the end of a project can bring in sufficient data for evaluation (especially in cases well covered from other sources, with data collections from other entities available for analysis, such as data from Eurostat, etc.); on other instances, a continuous data collection can be a must, especially when a course of development is to be mapped and the desire is to shed some light on the process and not only on the outcome.

### 3.5.2 Collaborative (Participatory) and Objective Evaluation

<sup>69</sup> Greene 1988, Mark and Shotland 1985 in Rossi, Lipsey, Freeman 2004: 51.

According to some authors<sup>69</sup> a **collaborative (participatory) evaluation is a team project of the evaluator and representatives of one or more stakeholder groups**. Stakeholders are directly involved in planning, conducting, and analyzing the evaluation with evaluator whose role can be ranging from a team leader to a resource person reached out to only when necessary.<sup>70</sup> The evaluator typically helps with more technical tasks, but the stakeholders have the right to decide as to what is to be measured and in what fashion.<sup>71</sup> However, the evaluator does not make judgments based on the data and does not give any recommendations.<sup>72</sup> Evaluator only urges stakeholders to reflect on the data and knowledge of the policy, but does not step into the process itself, leaving

<sup>70</sup> Rossi, Lipsey, Freeman 2004: 51.

<sup>71</sup> Ibid.

<sup>72</sup> Weiss 1998: 99.

- the assessment and recommendation creation fully in the hands of the stakeholders.<sup>73</sup> In contrast to the empowerment evaluation which aims at evaluation conducted by community members, the participatory (collaborative) evaluation should promote an equal share of control by all stakeholders.<sup>74</sup>
- <sup>73</sup> Ibid.
- <sup>74</sup> Ibid.
- <sup>75</sup> Shulha et al. 2015. Shulha et al.<sup>75</sup> look deeper into this type of evaluation, bringing in principles based on their own empirical research of collaborative approach in evaluation, shedding more light into the usefulness of this evaluation type in aspects such as: creating relationships among the stakeholders; developing a shared understanding of the policy; or promoting evaluative thinking.
- <sup>76</sup> Ibid. On the other hand, an objective evaluation can be defined. According to Patton<sup>76</sup> “*traditional social science research methods have called for objective, neutral, and detached observers to measure the results of experiments and studies.*” But professional evaluation, as Patton<sup>77</sup> points out, requires involvement of the stakeholders to assure utilization of evaluation’s results. Therefore, it must be decided to what extent the stakeholders are to be included.<sup>78</sup> **In order to achieve as objective evaluation approach as possible, the involvement of stakeholders must not only be transparent, but must have no influence on data analysis and interpretation whatsoever. Evaluators themselves also need to be aware of their biases and preconceptions, making any and all potential influences on their evaluation-related work transparent** (e.g. in case of conflict of interests, etc.). At the same time, use of objective measurements and overall well-established and reasoned-for methodology is necessary as well, as chapter 4 refers to in detail.
- <sup>77</sup> Ibid.
- <sup>78</sup> Ibid.

This distinction is easy to see in case of independent evaluation conducted by a professional evaluator or evaluating body, and empowerment or collaborative approaches in which the evaluator is more in a supportive role, enabling the stakeholders to take very active part in the whole process. Naturally, this distinction can be much less apparent in cases such as the typical internal or external evaluation which is conducted, to some extent, in collaboration with the stakeholders as well: one of the stakeholders is usually commissioning the evaluation itself, and many other stakeholders are usually in the role of informants, making the objective and participatory evaluation almost always a matter of concrete setup and context.

### 3.5.3 Goal-based and Goal-free Evaluation

- <sup>79</sup> Ibid. During the evaluation, the evaluator assesses the policy in question. As Patton<sup>79</sup> points out: “*When making determinations about the appropriateness, adequacy, quality, efficiency, or effectiveness of program operations and results, evaluators may rely on existing criteria provided in laws, regulations, mission statements, or grant applications. Goals may be clarified, and targets for performance may be given in such documentation. But in some cases, evaluators are not given such criteria, and may have to seek guidance from stakeholders, professional standards, or other evaluation studies to help them make judgments. When little available guidance is given, evaluators find themselves constructing the evaluation criteria – or operating in a goal-free state.*”

In other words, **goal-based evaluation is the typical evaluation based on objectives of the given policy, and making assessments based on these objectives**, in the simplest form stating whether the objectives have or have not been reached. **Goal-free evaluation is a rather atypical situation which either requires the evaluator to create policy goals based on expert knowledge or consultations, or to limit the evaluation to stating the evaluation findings, but not stepping up to assessing the findings against the policy objectives**, since these are not available. It is noteworthy that, should such case arise, this goal-free state in itself is one of important evaluation findings.

## 3.6 How to Choose an Appropriate Evaluation Type?

Various authors argue for different criteria when determining which evaluation type to use. Chen<sup>80</sup> suggests the following steps when looking for the appropriate type:

<sup>80</sup> Chen 2015: 51-53; modified by author of this publication.

- I. In cooperation with the stakeholders, **identify the concrete focus of the evaluation**: Is the evaluation to tackle the implementation of a policy? Is the evaluation to focus on impacts of the policy? At what stage of the policy implementation is the evaluation to take place? Be as specific as possible on the expectations of the stakeholders and facilitate the discussion in order to get concrete answers determining the aims of the evaluation as precisely as possible.
- II. **Identify potential audiences**. Is the evaluation to serve internal purposes, or is it meant for wide audiences outside of the circle of stakeholders? Should it serve both internal and external purposes? When answering these questions and combining it with the aims of evaluation identified above, certain evaluation types will become apparently more appropriate than others. E.g. evaluation produced for the purposes of communicating the results of a given policy to the wide public will require different approach and scope than an evaluation which is for the internal purposes only and aims at optimizing the implementation processes of the policy.
- III. **Communicate the details of the chosen evaluation type** and research method **back to the stakeholders**. This step enables all stakeholders to keep ownership of the whole process as well as to provide feedback and help the evaluator to optimize the evaluation strategy before commencing further evaluation processes.

<sup>81</sup> Frechtling et al. 2010:

Other authors<sup>81</sup> state that **choosing an evaluation type also depends on the type of questions the evaluation is to address**. This is in line with the steps described above in chapter 1, since identifying focus and audience for the evaluation is necessary to form evaluation questions. To form explicit questions may help to guide the whole process more precisely, keeping on track and not diverting into interesting, albeit not important details along the way.

All in all, it needs to be pointed out that the **evaluation types are only theoretical concepts** presenting the wide range of the potential evaluation processes which can be utilized. In practice, these types often merge and become hybrids in order to enable the evaluator to reach the goals and answer the evaluation questions.

As an example, evaluating youth policy towards rural areas in a given country can be a complex process involving internal evaluation in the first stage, focusing on the implementation processes of the policy (formative internal evaluation) and aiming at further development of the implementation in the future (e.g. after the first pilot phase of the policy implementation); this formative internal evaluation can also involve collaborative evaluation approaches in order to collect the data as well as to enhance the ownership of the given policy among the stakeholders who are involved in the policy implementation. Subsequently, this initial phase can be followed by an external evaluation focusing on outcomes (summative external evaluation) and aiming at presenting results of the policy in a certain time period (e.g. the first year of policy implementation) to the public. And this summative external evaluation can involve empowerment evaluation approaches in some concrete rural areas in order to collect the data as well as to increase awareness of the local residents of the policy in question.

This intentionally complicated and complex example sheds some light on the potential utilization of the evaluation types. **The evaluation types are to serve the evaluation purpose** and can be used to the fullest as a single approach as well as to be combined with other types or used only as a small-scale approach within a wider evaluation scheme.

In order to further broaden the knowledge of the aforementioned evaluation types and to link them with concrete processes to take place during the evaluation process itself, concrete social research methods which can be potentially useful in any of these types are listed and described below.

# 4.

# Outlining Evaluation Methodology

The next step in conducting evaluation is outlining a methodology. Methodology is a term which describes a set of social science methods and the way they are used in a particular study, including any evaluation endeavours. When thinking of the evaluation methodology, it is useful to keep in mind the whole process which is closely linked to the usage of the methodology in practice, as summed up below<sup>82</sup>:

<sup>82</sup> C.f. Frechtling et al. 2010: 15-31; modified by author of this publication.

- I. **Development of an evaluation design.** When developing an evaluation design, these are the main steps:
  - Determining what type of evaluation is required to answer the evaluation questions.
  - Selecting a methodological approach and concrete data collection instruments.
  - Selecting a research sample, i.e. the sum of individuals to be targeted during data collection phase. This procedure is commonly referred to as “sampling”, and puts together a group to be tested, therefore called a “test group”. This group needs to possess information relevant to the evaluated policy, usually directly or indirectly subjected to the policy in question.
  - Selecting a comparison group to provide a reference frame for the evaluation in question, if necessary. Such group of respondents possesses the same qualities as the research sample but is not subjected to the evaluated policy and is hence usually called “control group”. Not all evaluation designs need to include the comparison group, for more details please see chapter 4.1.3.3 covering experimental designs.
  - Determining timing, sequencing, and frequency of data collection. Data collection is a general term for using research instruments and gather information from respondents (research sample and comparison group). Planning data collection steps is vital in order to put it in line with other processes and to determine human and financial resources necessary to conduct the data collection exercise.
- II. **Collection of data.**
- III. **Analysis of data.**



Choosing appropriate research design is crucial, since it constitutes a roadmap leading from the evaluation questions to their answers. The choice is dependent on all of the aforementioned points, from the evaluation questions, through the availability of individuals to include into the evaluation (research sample, and if applicable also the comparison group) all the way to the data collection and analytical capacities available during the evaluation process. All of these aspects, when explicitly formulated during the preparatory phase of the evaluation, often provide fine guidance when it comes to choosing a particular research methodology to use.

The most common research methodologies used in evaluation are listed in this chapter, and briefly explained, accompanied by references for further reading. This provides a basic overview of the research methods to be used in evaluation either as a combination (e.g. individual interviews and several focus groups) or as a single approach study (e.g. questionnaire via a large-scale online survey). Two main groups of methodologies are outlined below: social research designs building broadly on research methods found in sociology, psychology, and other related social science fields; and specific evaluation designs which are used in specific evaluation contexts, e.g. in cost-related matters.

## 4.1 Social Research Designs

General social research designs often used in evaluations are presented in this sub chapter, providing for a list of the most common approaches. The basic division between the quantitative and qualitative research approaches is outlined, with concrete research methods listed under each of these broad approaches. Mixed methods are introduced as well, bringing in an interesting link between the quantitative and qualitative traditions.

### 4.1.1 Quantitative Approach

Quantitative approach stems from the natural science tradition of gathering evidence on reality which can be observed and understood in the same fashion by every individual. This philosophical tradition is called positivism. As an example, the length of an object can be observed and understood by using one system of measurement and unless the object changes, the length of the object will always be the same, no matter how many times we take the measurement or who takes the measurement.

This approach applied in social sciences is trying to quantify (hence the name) measured phenomena in order to allow the understanding of the phenomenon in question. Ideally, such measurements will bring the same results if repeated (as is the case in the length measurement example above) and will refer to the phenomenon in question.

In practical terms, real-life concepts are attributed numerical values in order to represent these concepts, thus making them measurable: a person can have a very healthy lifestyle (represented by number 10, for instance) or a very unhealthy one (represented by number 0, on the opposite end of the scale), or have a combination of good and bad habits, resulting in average lifestyle (and therefore be represented by, for example, number 5 in the middle of the aforementioned scale). By assigning a number to each individual in the research

sample on this particular scale, we can use statistical methods to process the data in order to see, e.g. what level of lifestyle there is in the group we are exploring: calculating an average will show the information rather clearly. Group average of above 5 suggests a rather healthy-living group of individuals, while average score below 5 suggest the opposite.

This example not only shows the basic logic of the quantitative approach in social sciences and evaluation, but also highlights two important traits of such approach: processing limited information about a large number of cases. In the example above, we only learn about the level of lifestyle in the whole group, but to know the reasons for being above- or over-average, further inquiry needs to be made. Therefore, the quantitative approach allows for generalization of the results, but may lack a deeper understanding of the reasons behind the results, while also suffering from the downside of only getting answers to the questions asked, i.e. respondents only fill in answers in a rigid format with no or limited options to elaborate further.

#### 4.1.1.1 Questionnaire

**Questionnaire is a tool used in quantitative approach methodology to gather large numbers of information. In essence, it is a written set of questions to which there are predefined answers for the respondents of the questionnaire to choose from and mark;** one individual fills in one questionnaire on the given topic. Once the respondent marks all their answers, these are transformed to numerical values (e.g. male is designated 2 and female 1) and transcribed into a database usually called a dataset. In the dataset, responses from a single person are all in one row with columns showing answers to the questions. In a dataset of 100 respondents and 10 questions, there will be 100 rows and 10 columns, answers for each person filled in by the designated numerical codes (e.g. all females are coded as 1 and all males as 2; all those who are very satisfied about an certain policy are coded as 5 and all those who are very dissatisfied are coded as 1, etc.).

Questionnaires can vary from a very short and widely distributed pieces usually called “opinion polls” to elaborated and very precisely targeted tools gathering information on a precisely defined population, labelled “surveys” (e.g. all users of a youth club between ages 15 and 18) in order to answer well-defined questions (e.g. to explore preferences of future development of the youth club agenda connected to mental well-being of young people). Questionnaire can exist in an online form, as a paper-based instrument, or even as a basis for an interview in which questions are asked to a person precisely as written down, and the answers are marked for the respondent. Target group needs to be taken into account when choosing the particular type of a questionnaire. In some cases, an online questionnaire may be suitable (e.g. all respondents have no trouble connecting to the internet), while in other cases a paper version can be a better fit (e.g. in cases where the internet connection can be a problem). There are cases in which an interview is the only way to gather responses for a questionnaire, e.g. in cases of respondents who cannot read as well as in cases of disabled young people with impaired vision.

Large numbers of responses (it is highly recommended to gather at least 100) subsequently allow statistical analyses which have a potential to generate findings applicable to general populations. Respondents should be chosen at random and the dataset ideally modelled (weighted) in order to mirror the general population as best as possible in basic characteristics, such as gender, age, or educational attainment. For example, trying to determine needs of young people in a certain rural region of a country does not necessarily require gathering answers from all of the young people in the given region. Instead, more than 100 responses are gathered from the young people in the given rural region, and these are subsequently analysed in order to provide an insight into

the general needs of the youth population in the area. Another “soup example” can be quoted: in case you need to determine the taste of a soup, you do not need to eat the whole pot; instead, you simply taste a spoonful, assuming that the taste will be the same in the whole pot. Using a questionnaire is basically the same approach – by using a sample population (the spoonful, i.e. the 100 young people who fill in the questionnaire), a generalization is made without the need to speak to all members of the general population (the whole pot, i.e. all young people from the whole rural region).

As in any other research method, it is important to acknowledge the advantages and disadvantages of the questionnaire use. Explicitly mentioning them helps deciding whether to include this research method in the evaluation process in question. As for advantages, the following can be mentioned:

- **Relatively low cost of data collection in comparison to other research methods.** Using a questionnaire can be rather straightforward, with the online version enabling to both collect the data and create a dataset as well as conduct basic analyses straight away, hence reducing the costs.
- **Potential for generalization and trend analysis.** In case the data possesses desired qualities, such as number of cases or representative profile, analyses results can be applied to all of the people in the general population, providing for a much-valued insights on large scale.

Disadvantages are summarized by the bullet-points below as well:

- **Expertise in creating a questionnaire is needed.** In order to create a functional questionnaire which brings in data enabling the evaluator to find answers to the evaluation questions, a tight link between the contents of the questionnaire and the evaluation questions needs to be established, requiring careful approach and expertise in utilizing particular questionnaire forms and using appropriate wording and stylistics.
- **Expertise in statistical data analysis is needed.** Although many online engines are capable of producing basic statistics, if more complex computations are required, a statistical analysis in a specialized software is necessary. In such case, the software as well as the statistical knowledge are a must and an expert might need to be hired.
- **Predefined answer choices provide limited information.** Simply, there are no answers apart from the ones provided in the questionnaire itself. In case a size of a village the young person lives in needs to be determined, it is hardly possible to ask an open question, since most people would struggle to give a precise number of inhabitants. Therefore, several options are created, such as: a secluded farm of a few people; a small village of up to 500 people; a village of up to 1000 people; a settlement of up to 5000 people; a large village of up to 10 000 people. This brings enough choices for the respondents to identify with; on the other hand, there are villages which are located in a close vicinity of a large city, effectively making such settlement much different from any other village in terms of frequency of public transports connections and many other important aspects. Without asking yet another question, such details can be lost and even provide baffling outcomes.
- **In-depth understanding of the analysis outcomes may require further research.** A statistical analysis of questionnaire data may suggest that young people from a given rural area are lacking opportunities to engage in organized youth work, although it is known that in terms of facilities as well as in terms of number of youth workers, this rural region is comparable to other regions with no such trouble. Is this due to lack of efforts of the local youth workers? Is it due to lack of experience in local youth centres? Is it due to negative attitudes of families in the particular region towards youth work as such? Questionnaire will likely not be helpful in answering such questions, since it is an unexpected result and the questionnaire was not designed to further follow this angle of inquiry.

- **Large numbers of responses are needed** in order to make the statistical analyses reliable. Since the whole idea of quantitative measurement in social sciences is based on statistical theory, and in particular on statistical inference (processes of deducing information from results of the research sample to the whole population, e.g. from the 100 young people to all of 1000 young people living in the same area), it is necessary to collect at least 100 responses from the target population (for details please see central limit theorem<sup>83</sup>). This does not mean that a survey of less than 100 responses is useless, but it saves much analytical work and provides greater reliability to the analyses to achieve collecting over 100 responses.

<sup>83</sup> De Vaus 2002:79.

All in all, it is obvious that while questionnaire is a useful instrument, it may be worthwhile to find an expert who knows how to put it into practice from designing through data collection all the way to the data analysis.

An example of a concrete questionnaire use in the youth policy evaluation can be found at the UNESCO website: 8th UNESCO Youth Forum Evaluation<sup>84</sup>. This evaluation was directly based on questionnaire data and provides an example of presenting survey results on a specific evaluation-related topic. Another example of a survey-based evaluation supporting “public governance arrangements for young men and women to engage in public life”<sup>85</sup> is Seven key findings from the Youth Governance Survey report by the OECD<sup>86</sup>. And an online large-scale multilingual questionnaire was also used in RAY study on monitoring of outcomes of youth mobility projects<sup>87</sup>.

<sup>84</sup> UNESCO 2014.

<sup>85</sup> OECD Governance 2019: 2.

<sup>86</sup> OECD Governance 2019.

<sup>87</sup> Fennes et al. 2011, 2012; Fennes, Chisholm 2013; Bammer et al. 2017; Böhler et al. 2019.

#### 4.1.1.2 Open Source Quantitative Data

In many cases, no original questionnaire needs to be designed and no survey exercise is necessary, since there are **databases of open source quantitative data available for analysis**. Examples of such are various indicators that regional, national and international bodies are collecting and providing, such as general Eurostat data on young people, specific Eurobarometer surveys on young people, or national statistical offices reporting on various areas of young people’s lives.

Utilizing all existent data should always be prioritized to collecting new information, since a new data collection always takes time of the informants and influences their willingness to participate in surveys in the future. With respect to this important consideration, it is advisable to only collect original data in cases no relevant data can be found. At the same time, data collection itself needs to be funded, hence using ready-to-use data in evaluations is a cost-effective way of covering, at least some, areas of interest.

Using open source quantitative data needs to take into account the original purpose of data collection: What research or evaluation questions were covered by the data? What was the general population? What was the sample of the survey? What reports are already available and could be potentially used to build upon? What is the correct way to quote the data source?

All of these questions, and more, need to be answered before using any open source quantitative data. Evaluator needs to be sure that they know the original purpose of the data collection and that this purpose fits the needs of the current evaluation in answering some of the current evaluation questions. It is also necessary to be explicit about the data being used and their original source: admitting and quoting the data back to the original author(s) is a legal and ethical must.

<sup>88</sup> Eesti Noorsootöö Keskus 2019.

<sup>89</sup> Tilburg University 2019.

<sup>90</sup> ESS ERIC 2019.

An example of such open source quantitative data is the Estonian Youth Monitor<sup>88</sup> which provides data on young people in a number of areas and with an opportunity to also explore trends over time. Further examples worth exploring, should there be a need for more detailed analyses are international studies, such as European Values Study<sup>89</sup> and European Social Survey<sup>90</sup>. Both initiatives provide open source datasets which can be further analysed to explore youth-related areas of relevance to the particular evaluation effort; both are also long-term projects providing data collected in various years, allowing for trend analysis as well.

#### 4.1.2 Qualitative Approach

Qualitative methods stem from a different philosophical tradition than the quantitative ones. While the quantitative approach is based on the presumption that the world around us can be measured and understood in a way common to all people, the qualitative approach builds on a premise that each individual has their own unique perception of reality and this perception contributes to the overall reality around us, since each of us acts on their own beliefs and mental images of reality. This philosophical viewpoint is called social constructivism: the overall reality is constructed in social interaction between individuals who act on their own perceptions of reality at the given moment. As a result of the abovementioned premise, the only reality that can be discovered is the one created at a particular point in time by actions and perceptions of people and their interactions.

To showcase an example, it is currently normal for youngsters of 16 years of age to vote in some countries such as Malta or Austria. A public debate in these countries reached a consensus of enabling the young people to have their say in the elections. When meeting with people from other European nations, their views on the legal voting age can differ with various arguments being presented. Most of these arguments would be the same for both the Maltese arguing for the lowering of the voting age to 16 and for anyone opposing it, but the conclusions would be different, since they are based on different realities. And, of course, there may be many people from all over Europe who appreciate the Maltese decision and wish for their countries to join the suit, but the current reality in their countries does not permit it to happen. These multiple realities not only of different countries, but of each individual being, are the basis of the social constructivism which poses that only by discovering the reality perceptions of the individuals in a given situation, a puzzle of the general reality can change. And again, if any single piece of this puzzle changes, the whole reality looks a bit different already.

<sup>91</sup> Goffman 1956.

An inspiring reading providing insights into the social constructivist view of reality can be found in Goffman's<sup>91</sup> work on roles individuals play in the society. While using a metaphor of a theatre, he explores various roles people represent in everyday life, showcasing the differences in behaviour of individuals across various contexts, and searching for an interpretation framework for human actions. The metaphor of "wearing masks" is well known. Goffman argues that each individual possesses a wide variety of "masks" which are used in different contexts: a mother at home, a supervisor at work, a citizen on the street, etc.

To illustrate the qualitative approach in evaluation further, an example from a simple youth project evaluation can be shown. While a questionnaire survey might identify the general level of fulfilment of expectations to be rather

average, qualitative methodology may come with more detailed pictures. Some youngsters may seem to be much excited about the project and find it matching their initial expectations, while others may be highly disappointed when it comes to their own expectations. This picture is more complex and less clear in the overall result than the one painted by the survey results (it is very difficult to deduce from a limited number of single cases to a general condition), but at the same time it can provide additional information leading to a deeper understanding than just a general level of fulfilment of the expectations: Is it because some of the youngsters were better informed of the contents of the programme than others and formed more realistic expectations? Was there a concrete aspect of a programme which made some youngsters lose attention, therefore leading to their expectations being left unfulfilled? Was the communication channel used in the project (academic language, mother tongue, foreign language, etc.) preventing some youngsters from fulfilling their expectations? All of these are important questions which may not be covered by the survey but are easily asked in an interview.

In practical terms, qualitative approach is typical by collecting and analysing detailed information about a smaller amount of cases or responses. In other words, by using face to face formats of data collection as well as by collecting various additional information (e.g. documents created by respondents, observation notes on particular activities, etc.), and in-depth understanding of the evaluated phenomena is created. The main medium of information collection is words and looking for individual meanings and interpretations of the respondents. Evaluation using qualitative approach is seeking to achieve understanding of the evaluated phenomena through the eyes of the respondents.

In the following subchapters, interview, focus group, observation and case study designs are outlined in order to provide a deeper understanding of the concrete tools provided by the qualitative approach in evaluation.

#### 4.1.2.1 Interview

**Interview is a one-on-one talk between an interviewee** (the person being interviewed) **and an interviewer** (the person leading the conversation and collecting data) in which a subject of interest is debated either freely (an open interview) or in line with a predefined script, usually called “interview guidelines” (semi-structured or structured interviews). This discussion usually takes place in person in a calm environment, nevertheless, other means can be used as well, such as online interviews, or phone interviews. In general, the interviews taking place in person are much better in establishing a rapport between the evaluator and the interviewee, while the alternative forms utilizing modern communication technologies can keep the time demands at the bare minimum, eliminating the need for travels.

Unlike in case of questionnaires, the choice of respondents for the interviews must be intentional, with a strong reasoning for choosing each of the interviewees. The chosen individuals should have a first-hand experience with the evaluated policy, for example: members of the target group of the policy; key stakeholders; administrators of the policy, and many others. Interviews can be divided into several subgroups, with different interview guidelines for each of the specific groups, as the questions asked need to fit the profile of the interviewee: it makes no sense to ask questions the interviewee does not know answers to.

Since interview is a lively talk between the interviewer and the interviewee, it is necessary to agree upon the method of the data collection. The easiest way is to record the interview, the interviewee needs to be informed of that beforehand and needs to provide consent with the recording. In case the recording is not possible (i.e. the interviewee does not provide the consent), making notes

is another, albeit much less comfortable, option. A recording is subsequently either listened to and analysed as a recording itself, or transcribed and analysed as a text, which provides the advantage of easily going back and forth as well as easily retracting direct quotations if needed. Last but not least, all data obtained via an interview need to be anonymized in order to prevent anyone connecting the findings and the interviewees. Generally, this is achieved using a combination of two methods: keeping the interview list confidential and only for as long as it is useful to the evaluator and then discarding it in a safe way; and never using any real names or any other data which could identify the person in any reports or publicly available materials (e.g. erase all mentions of geographical locations, names of institutions, and any other information potentially linking the interviewees to the collected data).

As is the case with the quantitative approach, there are advantages and disadvantages to the qualitative approach as well. Advantages can be summarized as follows:

- **Qualitative methods have a potential to provide large amounts of detailed data.** Even though qualitative methods also need to be prepared in terms of interview guidelines or observation templates, there is always a potential for flexibility. Asking additional questions or adding more information to the observation template to create as complete a picture as possible can easily be done. This provides the evaluator with an opportunity to follow up on unexpected or new information and fill in blank spots on the go.
- **It is not necessary to have large numbers of respondents.** Unlike in the quantitative approach, generalization is not the aim of the qualitative inquiries. Therefore, even a relatively low number of well-chosen key respondents can provide valuable and deep insights into the evaluated matter, and there is no minimum number of respondents that need to be approached. As a rule of thumb, it is advisable to approach such a number of respondents or collect as much data as is possible to utilize in data analysis; in case analytical capacities are not limited, then data collection continues until information gathered through the research instruments becomes repetitive.

Disadvantages can be summed up as follows:

- **Data collection and analysis is relatively costly and time demanding.** Since the data collection is dependent on personal engagement in interviews and observations, the costs and time allowance need to be planned for. Each of the respondents needs to be approached in advance, and a meeting scheduled, which often takes up more time than expected. All of the collected data needs to be analysed, i.e. all interviews need to be listened to again in detail (or read through, in case a transcript is available), and all the documents need to be studied as well. All of these activities take time and need to be budgeted for.
- **Generalization is not the aim of the qualitative approach.** While the quantitative approach aims at getting the overall picture, it is the goal of the qualitative approach to provide detailed insight into the perceptions of the actors. This does not mean that the findings of the qualitative evaluation are useless in defining the overall picture, but it needs to be treated carefully and linked always to the source from which the information came.

#### 4.1.2.2 Focus Group

**Focus group is an interactive discussion led by a moderator** and has a potential to capture perceptions, experiences, beliefs or expertise of a small group of people about a given topic, project, or program.<sup>92</sup> Importantly, focus groups should bring together people of a similar background and in a similar relation to the evaluated policy in order to boost the discussions and use the group dynamics to getting as much data on the topic as possible. If the evaluation identifies more than one homogenous group of actors, it is possible to conduct multiple focus groups, exploring different angles with each group.

<sup>92</sup> Chen 2015:158.

For example, it is possible to conduct a focus group with handicapped young people in order to see their views of the latest changes in the health-care policy. At the same time, it might be beneficial to conduct a focus group with doctors focusing on the given age group, to hear their opinions as well. And of course, it might be potentially useful to hold a focus group with parents, personal assistants or other people in a similar position, i.e. in touch with both the young handicapped people and with the doctors and health-care services. Each of these focus groups might have a different guideline and will provide a unique point of view, bringing an opportunity to compile as detailed and full picture as possible.

#### 4.1.2.3 Observation

When conducting an **on-site observation**, the evaluator in person overlooks implementation processes of given policy measures in order to describe major aspects of such measures, identifies strong and weak suits and analyses their causes.<sup>93</sup>

<sup>93</sup> Cf. Chen 2015:158.

Based on the level of engagement, the participation can be participatory, where the evaluator for many intents and purposes acts as a recipient of the youth policy measures, i.e. acts as a member of the target group of a given policy. On the other hand, the observation can also be completely non-participatory, resulting in a situation in which the evaluator stands aside and acts as an impartial observer, observing only without any direct participation in any of the processes. At the same time, there are many in-between roles for the evaluator to take, combining the participatory and non-participatory approaches to the best outcome in the given context.

In order to conduct a successful observation, the evaluator needs to prepare beforehand by studying all available documents on the processes in question, and deciding whether a structured observation or an open, unstructured one, suit the case. In a structured observation, a grid for field notes is created beforehand, hence explicitly marking the focus of an observation for the evaluator, using, e.g. phases of the process as categories to be observed, or transversal topics to be focused upon. This helps the evaluator to focus on certain aspects of the overall process, while also limiting the focus to situations outside of this grid. In the unstructured observation, the evaluator takes field notes simply based on the momentary situation, keeping their mind open to all and any aspects occurring in the observed processes. Each of these approaches has its advantages and downsides and often, again, combining them may provide for an effective tool: having a general observation grid ready, but also keeping space in the field notes for any other observations.

As mentioned above, the main format of the data collected during an observation is often field notes, since video recording of the observation is hardly ever possible<sup>94</sup>. Such notes, based either on the structured approach or any form of less-structured observation, are analysed in order to reveal details not available via other forms of data collection. Combining observation and other data sources can provide useful insights into the perceptions of the processes and policies in question, focusing on discrepancies between the contents of the observation field notes and of information provided from other sources. Needless to mention that being present when some of the policy implementation processes are taking place, interviewing actors present on the spot can prove to be an efficient use of time and human resources.<sup>95</sup>

<sup>94</sup> Explicit consent with the video recording would need to be obtained from all involved personnel, including the consent from legal guardians in case of minors. At the same time, the act of recording itself may affect the way people behave, and therefore disturb the results of the study.

<sup>95</sup> Cf. Chen 2015:158.



When exploring outcomes of a youth policy focusing on youth engagement, it might be useful to engage in observations of some key implementation events, such as community planning events, youth participation workshops, youth initiatives funded via specific dedicated grant schemes, etc. Besides using the time at the events for observation purposes, it can well serve as a gateway to important respondents. Taking time and speaking, informally, with the young people themselves, may open doors to future interviews, and shed more light on the proceedings from the perspective of the target group members. At the same time, it is important to observe missing components as well: Are there any subgroups of young people completely missing at these events? Do other actors, such as local policymakers, take part? Are all components of the event taking place as planned?

#### **4.1.3 Mixed Method Approach**

Combination of quantitative and qualitative methods can provide for a more detailed picture than using any of them separately. Mixed method approach is not, however, characterized only by employing both quantitative and qualitative methods within one evaluation scheme, but requires the methods to be combined in order to complement each other. The most common approach is to employ either qualitative or quantitative methodology first and based on the findings brought by this particular approach, the other methodology is used, further exploring and deepening the findings.

As an example, focus groups with young people can be used to initially explore a new approach to youth spaces in a given geographical area (a village, a suburb, etc.), and based on the findings stemming from the analysis of the focus group data, a questionnaire is designed which is subsequently distributed to a large population of young people, determining to what extent the initial focus group findings are applicable on a larger scale. That way, existent and desirable as well as missing spaces for young people can be discovered. The questionnaire is already building on the findings from the focus groups, enabling it to be much better directed than if only based on theoretical literature or policy documents.

Apparently, multiple data sources and data collection mechanisms are a defining feature of this evaluation approach, aiming at strengthening the evaluation outcomes by using complementary methods to the advantage of the whole evaluation process. Case study, action research and experiment are mentioned as rather well-established representatives of the mixed method approach in evaluation. Nevertheless, the mixed method approach is typical by combining various research methods ad hoc in order to best fit the peculiarities of each evaluation context and objective, and therefore any mixture of research methods which is used in an interplay (i.e. findings from one phase of the evaluation feed into the next) can be considered a mixed method approach.

##### **4.1.3.1 Case Study**

Describing an overall characteristic of a given policy through an evaluation can be achieved by collecting general or concrete data and deducting or inducing to generate evaluation findings, as is the case in the abovementioned research and data collection techniques. A different approach, however, can also be used: by choosing appropriate examples (cases) and describing them in detail, an evaluation can demonstrate the basic characteristics of a given policy on real-life situations.

Case study, as suggested above, is a process of choosing a “case”, i.e. the target of the study, and subsequently use as many data sources as possible in order to describe the chosen case in as much detail as possible. In doing so, and provided that the case has been chosen well, valuable information can be collected, and colourful findings may be drawn, shedding new light not only on the case in question, but also on the whole evaluated policy.

Choosing a particular case is the most vital decision when engaging in a case study. Such a case must be selected so that by exploring it, either overall policy in question or its important elements are looked into. Choice of cases varies from particular segments of a policy implementation process (e.g. concrete implementation measures), through concrete organizations (e.g. local youth initiatives and NGOs using a particular policy to their advantage), and even individuals (e.g. a young person in need of long-term health treatment and hence engaging in a particular policy over long periods of time). Since this is best described via examples, two potential case studies are shown below.

In the first example, policy aiming at developing digital youth work in rural areas of a country is evaluated. Based on the initial assessment by the evaluator, one of the rural areas is selected as a case in a case study, since preliminary reports show that implementation of the policy in question brings promising first results in this particular geographical area. In this example, the case represents all aspects of policy implementation, but limited to a smaller area, and hence available for a detailed study. The evaluator would travel into the chosen rural area to conduct interviews; go over any relevant documents related to that particular area and policy in question; observe the implementation process; and even utilize any and all quantitative data available and related to the situation of the young people in the given area. That way, details of the policy implementation as well as information on the successful efforts and obstacles along the way can be explored, and potentially linked to the overall policy process in other rural areas of the given country.

In the second example, the same policy is to be evaluated, but a different case is chosen. Instead of going into one particular rural area (geographical location), the evaluator chooses one particular measure to be the examined case: funding programme supporting education of youth leaders in rural areas in digital youth work. This case is chosen for several reasons: it is newly introduced, aims at all rural areas in question, and generates substantial funding opportunity for the local youth leaders. By choosing this particular case, the evaluator aims at gathering information on its implementation across all of the targeted rural regions, describing its implementation in detail. Such results can be useful when discovering implementation mechanisms which are successful as well as those which are hindering achievement of the policy objectives. That way, critical implementation failures on the local levels as well as misunderstandings between the national and local level actors can be explored, and examples of good practice avoiding these pitfalls can be identified. All of the above-mentioned information is useful in the overall policy evaluation.

As seen above, case study can be a powerful evaluation tool, provided that the case which is to be explored is chosen well. Importantly, further utilization potential of the case study findings within the evaluation framework needs to be taken into account. In other words, the case study needs to fit into the general evaluation process by providing such information which well complement all other research methods employed in the given evaluation. At the same time, the case needs to have a potential to provide rich data from various sources in order to allow the evaluator to complete a full and detailed description of the case itself.

#### 4.1.3.2 Action Research

Direct and immediate utilization of research and evaluation findings which feed back into the process of policy implementation and lead to adjustments in the policy implementation process itself is the basic scheme of the action research approach. Similarly to the case study, action research also utilizes multiple data sources in order to obtain as much information on the policy in question as possible. The close link described above between the evaluation findings and adjustments of the policy implementation process is a basic step which can be repeated as many times as is desirable in order to optimise the policy delivery and increase the policy achievement levels. The evaluator is accompanying the whole process, collecting data, analysing them, providing with findings and recommendations, overseeing and providing expertise on the policy adjustments, and again collecting data on the adjusted policy implementation processes.

Even though this process may seem complex and therefore hardly useful in youth policy evaluation, it is used on a regular basis when quantitative data are collected on various aspects of policy implementation, and based on the outcomes of the analyses, the implementation is adjusted.

Youth policy implementation on the local level can provide examples of action research in youth policy evaluation. Youth organizations implementing or contributing to an implementation of various youth policies can repeatedly perform (self-) evaluations, identifying areas to be improved, and checking again in the future. Utilizing interviews with the youth leaders, surveys of the young people, and observations on the spot, such action research can lead to substantial improvements in the youth organization performance, and hence also in the delivery of youth policies on the local level.

#### 4.1.3.3 Experiment and Quasi-experiment

Experiment is an artificially created situation planned and controlled by the researcher in order to observe and examine behaviour of the experiment participants. Experiment consists of dividing participants randomly into two groups – one is exposed to the policy (test group) and the other is not (control group).<sup>96</sup> Subsequently, any differences in qualities of the groups at the end of the policy implementation are assumed to be related to the policy outcomes.<sup>97</sup> All in all, the basic premise of the experimental design is that dividing participants randomly into two groups, hence having two qualitatively comparable groups of people (e.g. the same number of university graduates, the same number of young people not in education, employment, or training, etc.) and allowing one group to be influenced by a given policy while the other group is not a recipient of the policy measures, leads to a situation where all and any differences in these groups should be due to the influence of the policy in question.

Setting up an experiment, apparently, is a very demanding endeavour and in many contexts the experiment is not feasible at all. In real-life situations, many variables (influences) cannot be controlled at all, while still being present (young people are often engaging in many activities outside of the youth policy framework) and effectively negating the basic premise of relating all differences between the test and control groups to the influence of the policy in question.

For example, an increase in levels of political participation in a group of young people targeted by the policy measures may well be due to other influences, such as grassroots activism present among the young people, not due to the successful implementation of the policy itself.

<sup>96</sup> Weiss 1998: 215.

<sup>97</sup> Weiss 1998: 215.

This type of research design also usually brings up ethical questions which may be yet another reason against conducting an experiment in evaluation.

It is hard to imagine providing one group of young people with a set of policy measures, while on purpose neglecting another only to create experimental situation for the purposes of policy evaluation.

Some answers to the downsides of the experimental design can be found in the quasi-experimental approach. Quasi experiments are based on the same premise as experiments (i.e. comparison of a group of people targeted by a given policy and a group of people not influenced by it), but instead of randomly and purposefully choosing participants for each group, a real-life situation resembling the experimental design is utilized.

That way, for instance, young people who participated in youth mobility projects, can be compared to those who never had such experience.

Provided the groups are rather similar in basic qualities (e.g. age, education, etc.), the basic premise again states that the differences detected between the groups are due to the influence of the youth mobility experience, rather than due to other influences or random differences<sup>98</sup>. Quasi-experiment is therefore much more feasible than the experiment as such, but the danger of misinterpreting the levels or origins of an influence of a given policy need to be considered very carefully, because there is often little or no control over the composition of the groups in question or other influences apart from the policy implementation processes<sup>99</sup>.

<sup>98</sup> Bárta et al. 2019.

<sup>99</sup> Cf. Chen 2015: 256.

## 4.2 Specific Evaluation Designs

Research designs specific to the evaluation domain are listed below. These are tackled separate, since they are not social research research designs as such, but rather consist of specific usage of research methods and techniques in a given evaluation context.

### 4.2.1 Needs Assessment

Needs assessment is a systematic process of identifying, measuring, and prioritizing the needs of a given target group, such as young people in a certain region in case of a regional youth policy. The aim is to provide knowledge basis for decisions concerning the target group in question by providing an overview of the existing needs. To reach this goal, research designs such as surveys, focus groups, interviews, secondary analyses of existing data, or a combination of methods, is used.<sup>100</sup> Needs assessment is well suited for ex ante evaluations, since it can provide useful input for a debate on future policy implementation setup (for details please see subchapter 3.4.1).

<sup>100</sup> Chen 2015: 101-102.

## 4.2.2 Efficiency Assessment

<sup>101</sup> Rossi, Lipsey, Freeman 2004: 60.

Efficiency assessment considers the relation between policy costs and its benefits.<sup>101</sup> It can be conducted utilizing various cost-related analyses listed below. Engaging in efficiency assessments “can be tricky and arguable because it requires making assumptions about the dollar [monetary] value of program-related [policy-related] activities and, sometimes, imputing monetary value to a program’s benefits”<sup>102</sup>. This warning applies also for readers and stakeholders who are critically assessing evaluation reports based on such approaches. Each efficiency assessment methodology needs to be reviewed in detail in order to determine reliability of any evaluation-related claims.

<sup>102</sup> Rossi, Lipsey, Freeman 2004: 60; square brackets added by the author of this publication.

### 4.2.2.1 Cost-Effectiveness and Cost-Minimalization Analyses

Cost-effectiveness analyzes a policy from the perspective of achieving given intervention outcomes in relationship to the costs as well as considering usage of different policy implementation strategies and determining their respective cost-effectiveness. Cost-minimization analyses attempt to find the least expensive way to fulfil the defined policy outcome.<sup>103</sup> Apparently, both of these methods aim at determining monetary value of either the policy implementation used in reality or an approach which can or might have been used for minimal costs.

<sup>103</sup> Rossi, Lipsey, Freeman 2004: 63; Weiss 1998: 244.

As an example, in order to achieve higher social cohesion in Europe, a mobility programme for young people can be set up, with a certain annual budget. At the same time, an annual budget can be dedicated to a set of measures in formal schooling, such as teacher training in supporting social cohesion in young people, and time allocations for various activities in schools. These measures can be compared in order to see what outcomes there are for each of them, and what costs are related to them. Recommendations can be made, based on the results, as to which of these measures should be funded in the future in order to get to the best result using minimal funding.


### 4.2.2.2 Cost-Benefit and Cost-Utility Analyses

Cost-benefit analysis translates costs and outcomes into the same unit of analysis (typically monetary) so they can be easily compared and analysed. Cost-utility analyses compare costs to utility which means the worth of a specific outcome for an individual or for society.<sup>104</sup>

<sup>104</sup> Rossi, Lipsey, Freeman 2004: 63; Weiss 1998: 244.

Using these approaches in assessing youth policy aiming at providing equal access to information for young people in rural areas, different methods of information delivery are explored in monetary terms: How much does it cost to deliver one piece of information to one person using (a) internet webpage (b) direct emailing (c) direct mailing of hard copies or (d) publishing articles in printed magazines? By determining the costs for each of those options (or any other options used or considered for use), and in combination with other data (e.g. how many people were reached at all using the different approaches, and how much was therefore the real cost of reaching out to a single young person), an evaluation of cost-utility levels of different approaches can be presented.

# 5. Debating Evaluation Ethics



Ethics are an indispensable part of an evaluation process and need to be considered from the very beginning of the evaluation endeavour but especially when choosing and using research methods, dealing with stakeholders, and when reporting results. The evaluation ethics are discussed in this chapter.

<sup>105</sup> Chen 2015: 16, Rossi, Lipsey, Freeman 2004:405.

<sup>106</sup> The American Evaluators Association 2004 in Chen 2015: 16-17.

Some authors<sup>105</sup> propose promoting professional ethics when conducting evaluation, specifically ethical principles of various evaluation bodies in the given context. Even though other approaches can be used, for the sake of outlining the main ethical dimensions of an evaluation process, an example of the ethical principles of the American Evaluators Association<sup>106</sup> are listed below:

- **Systematic inquiry.** Evaluators conduct systematic, data-based inquiries.
- **Competence.** Evaluators are competent to perform tasks within the evaluation process.
- **Integrity and honesty.** Evaluators ensure honesty and integrity of the entire evaluation process.
- **Respect for people.** Evaluators respect the security, dignity, and self-worth of the respondents, program participants, clients, and other stakeholders.
- **Responsibilities for general and public welfare.** Evaluators articulate and take into account the diversity and values that may be related to the general and public welfare.

**Systematic inquiry emphasizes the need for objectivity**, which is and needs to be an integral part of any evaluation process and is stated in the definition of evaluation itself. Using social science research methods is one of the ways to ensure the approach indeed is systematic and based on empirical data, however, data analysis needs to be conducted in a competent and honest manner to ensure the empirical evidence is treated in an ethical way and the final picture is as true to the evidence gathered during the evaluation process, as possible.

**Competence of evaluators is vital throughout the whole evaluation process.** Conducting evaluation in line with all research-related expectations as well as compliant with general evaluation procedures, following the evaluation questions while navigating the complex terrain of stakeholder ties without endangering confidential information from one stakeholder in relation to the others requires competent and systematic approach from the beginning to the end of the whole process. Again, an example may be taken from the Guiding Principles for Evaluators<sup>107</sup> of the American Evaluation Association:

<sup>107</sup> American Evaluation Association 1995 in Weiss 1998: 94-95.

- 1. Evaluators should possess the education, abilities, skills, and experience appropriate to undertake the tasks proposed in the evaluation.**
- 2. Evaluators should practice within the limits of their professional training and competence and should decline to conduct evaluations that fall substantially outside of their expertise.** When accepting conducting an evaluation, evaluators should make any significant limitations clear to the commissioning actor, including any limitations connected to the expert profile of the evaluators themselves.
- 3. Evaluators should continually seek to maintain and improve their competencies,** in order to provide the highest level of performance in their evaluations. This continuing professional development might include formal coursework and workshops, self-study, evaluations of one's own practice, and working with other evaluators to learn from their skills and expertise.

To act with the **honesty and integrity** means for the evaluator to be open and truthful with respondents and to treat them with respect, candour and honesty.<sup>108</sup> This means to provide the respondents with enough correct information about the study to enable them to decide if they are willing to answer the questions and permit observations, including the understanding about the use of the provided information within the wider evaluation scheme.<sup>109</sup> To achieve this understanding between the evaluator and the informants and respondents, an informed consent is used.

<sup>108</sup> Weiss 1998: 93.

<sup>109</sup> Weiss 1998: 93.

**Informed consent** refers to a document including all necessary information on the given evaluation to enable the potential respondent to decide on their participation without being tricked or coerced into it.<sup>110</sup> Although it is necessary to adjust any informed consent to (a) the legal environment the evaluator operates at, both national and international, taking into account all relevant legal obligations (e.g. the GDPR), and (b) the particularities of the evaluation itself, such document usually includes sections listed below:

<sup>110</sup> Weiss 1998: 93.

- **Title of the evaluation.** What is the preliminary name of the evaluation report? What is the name of the evaluated policy?
- **Details of the evaluator.** Who and on behalf of what entity is conducting this particular evaluation?
- **Details of the commissioning institution.** Who or what institution is the assigning and contracting entity in this evaluation? Who asked for the evaluation to be conducted?
- **Purpose of the evaluation.** What is the evaluation trying to achieve? What are the main evaluation questions? What audience is the evaluation conducted for? What publishing of results is foreseen and with what outreach?
- **Procedures to be undertaken by the informant.** What is required of the informant? What information and in what form is to be collected? How long will it take?
- **Confidentiality clause.** How is the anonymity of the respondent ensured? What measures are in place to secure the anonymity? What data, on the other hand, may be published? What further operations will be performed on the data? Is any sharing of the data foreseen?

- **Information on the voluntary nature of the participation.** Is the informant deciding on their own to take part in the evaluation?
- **Declaration of consent.** Does the informant agree with all of the aforementioned information?
- **Dates and signatures.** Do both the evaluator and the informant come to an agreement to proceed?

One of the critical ethical issues mentioned in the context of the informed consent is **confidentiality and anonymity**. By and large, all information collected during the research need to be held in strict confidentiality and in line with the contents of the signed informed consent.<sup>111</sup> In most cases and unless there are strong reasons for deciding otherwise, only the evaluation team should have access to the information about informants and respondents as well as to any raw data, i.e. any material directly collected from the informants, such as interview recordings or filled-in questionnaires.<sup>112</sup> Reporting of any information can only be done anonymously and it strictly must not be possible to identify any of the respondents. To achieve these goals, data are used in complex summaries and in case direct quotes from respondents are necessary in the report, the identity of the speaker is carefully covered using alias names.<sup>113</sup> In line with the anonymization of the data, all personal information need to be removed from documents with interviews, from questionnaire datasets, and field notes and replaced with codes or numbers where necessary.<sup>114</sup>

<sup>111</sup> Weiss 1998: 93.

<sup>112</sup> Weiss 1998: 93-94.

<sup>113</sup> Ibid.

<sup>114</sup> Ibid.

Another of the ethical issues is also the problem of **reciprocity**. When conducting an evaluation, an evaluator asks for informants' time, knowledge and information, usually without giving anything concrete back.<sup>115</sup> What can constitute an exception are the occasions in which financial or other incentives are foreseen for the informants in order to compensate for the time they devote to the evaluation efforts in providing information. If this is the case, it should also be listed in the informed consent document and taken into account when tackling the reciprocity issue. What is seen as an ethical obligation in any case is providing study results to the people who contributed with information to the study.<sup>116</sup> Weiss<sup>117</sup> adds that "*respondents who receive an accounting of what the study found out will often be interested in seeing that members of the decision-making community pay attention to the results*", which can constitute a positive impact in initiating or facilitating a public debate on policies in question. Reciprocity should always be openly tackled from the very beginning of the evaluation, since it can have budgetary implications, provided that the evaluator wishes to reimburse informants for their participation in the evaluation process, and it needs to be taken into account when deciding on the format and processing of final reporting of the evaluation itself.

<sup>115</sup> Weiss 1998: 95.

<sup>116</sup> Ibid.

<sup>117</sup> Ibid.

The evaluator has a responsibility for reporting the results of the study fully and honestly but this can be in conflict with the **protection of the interest** of stakeholders connected to the policy.<sup>118</sup> The evaluator needs to be sensitive to stakeholder needs and interests, but keep in mind that the obligation to the honesty and integrity of evaluator's work is always paramount.<sup>119</sup> The American Evaluation Association's Guiding Principles<sup>120</sup> can again provide further insight into this ethical issue:

<sup>118</sup> Weiss 1998: 109-110.

<sup>119</sup> Ibid.

<sup>120</sup> American Evaluation Association 1994 in Weiss 1998: 109-110.

*"Because justified negative or critical conclusions from an evaluation must be explicitly stated, evaluations sometimes produce results that harm client or stakeholder interests. Under this circumstance, evaluators should seek to maximize the benefits and reduce any unnecessary harms that might occur, provided this will not compromise the integrity of the evaluation findings. [...] Where possible, these issues should be anticipated during the negotiation of the evaluation".* In other words, the commissioning institution should be openly informed of the necessity to stick to the same agreed standards for both positive and negative findings of the evaluation. In case a public report is agreed upon, the agreement should be explicit in not interfering with the evaluation findings on one hand and producing balanced and objective findings on the other.



**Ownership** of the data collected during the evaluation constitutes another ethical issue. The evaluation process comprises of collecting data as well as of preparing evaluation reports, and it is necessary to have explicit agreements on the ownership of both the primary data (information collected from the informants during the evaluation) and of the evaluation reports.<sup>121</sup> Weiss recommends handle this issue in early phases of the study, with both the evaluator and the commissioning institution to agree upon guidelines concerning ownership and release of the data and reports.<sup>122</sup> This is necessary also with respect to the informed consent document described above: the ownership and release of the data needs to be shared with the informants for their consideration.

<sup>121</sup> Cf. Weiss 1998: 292.

<sup>122</sup> *Ibid.*

As shown in this chapter, ethical dimensions of the evaluation process can be complex and are best dealt with openly, explicitly, and early in the evaluation process. Some of the ethical considerations relate to the evaluation reporting, debated in detail in the following chapter.

# 6. Reporting on Evaluation



<sup>123</sup> Rossi, Lipsey, Freeman  
2004: 377.

**Reporting on evaluation comprises of activities through which evaluation findings are made available to relevant audiences.**<sup>123</sup> This being said, these audiences are to be identified at the beginning of the whole evaluation enterprise and agreed upon with the commissioning institution of the evaluation. At the same time, number of reports, their scope, requirements, and confidentiality of each of them is to be debated and agreed upon. To some commissioning institutions, a single all-encompassing public report may be enough; others might require a radically different approach, with the full report only available to them and specific reports published for different audiences and in different formats.

<sup>124</sup> Rossi, Lipsey, Freeman 2004: 377.

<sup>125</sup> Rossi, Lipsey, Freeman 2004: 377-380.

As some authors<sup>124</sup> emphasize, the reporting is a critical responsibility of all evaluators, and it is the reporting which usually brings the whole evaluation to both fruition and closure. To reach maximum usefulness, the evaluation findings need to be communicated in an accessible and intelligible manner in order to be available to various stakeholders.<sup>125</sup> In order to reach this ambitious goal, evaluation reports typically contain of six major parts:

- **Evaluation background.** Reasons for commissioning of the evaluation, stakeholder identification and description, conflicts of interests identified and mitigated, and the background of the evaluator or the evaluating entity are all clearly stated.
- **Evaluation questions.** The main evaluation question and all specific questions are presented.
- **Evaluation methodology.** In connection to the previous two sections, the methodology of the evaluation enterprise is explained.
- **Evaluation data.** Process of data collection and data analysis is presented.
- **Evaluation findings.** Findings based on the empirical data and interconnected with the context and reality of the given policy are presented.
- **Evaluation conclusions (and recommendations).** Final assessment based on the evaluation findings section, clearly structured and linked to the evaluation questions, is shown. Recommendations for further development of the policy are usually also part of the report, unless explicitly agreed otherwise with the commissioning institution.<sup>126</sup>

<sup>126</sup> Frechtling et al. 2010: 44; modified by author of this publication.

The aforementioned sections only represent a general outline and as every evaluation is different, every report needs to be tailor-made in line with the evaluation setup, including the agreements with both the stakeholders and commissioners of the whole process. Single sections can also be in a different order, for example Weiss<sup>127</sup> recommends opening the whole report with the findings section to sum up the main points. In any case, an accurate part describing methods of the evaluation is an integral part of every final evaluation report.<sup>128</sup> It is advisable to also devote time to secondary reporting. This may include abbreviated version of full primary reports; special reports presented in attractive and accessible formats; oral reports complete with slides at various conferences; and if appropriate, also multimedia outputs (e.g. videos, podcasts, etc.) to be presented online.<sup>129</sup>

<sup>127</sup> Weiss 1998: 295.

<sup>128</sup> Rossi, Lipsey, Freeman 2004: 380.

<sup>129</sup> Cf. Rossi, Lipsey, Freeman 2004: 381.

Although the conflict between the interests of the evaluation commissioning institution and the evaluation findings is tackled in the ethics-related chapter 5 above, it is worthwhile to mention a specific strategy of dealing with disagreeing positions in the evaluation report. First and foremost, it is important to listen to any comments of a commissioning institution or stakeholders with an open mind since they can provide interesting insights or interpretations.<sup>130</sup> At the same time, in case the suggested changes would disrupt the integrity of the report, these need to be rejected and an offer can be made to add a specific appendix which summarizes the objections of the commissioning institution.<sup>131</sup> Weiss<sup>132</sup> rightly underlines that *“a major contribution of evaluation is to provide an independent assessment, so that informed publics can trust the findings.”* This ethical consideration is paramount and can be achieved via the aforementioned reporting strategy without neglecting objections from the commissioning institutions or stakeholders.

<sup>130</sup> Weiss 1998: 304-305.

<sup>131</sup> Ibid.

<sup>132</sup> Weiss 1998: 304.



# Conclusions

Evaluation can be sometimes perceived as a straightforward research exercise with uncertain outcomes conducted by expert outsiders with no engagement in the given policy area, and as such is seen as a troublesome necessity by some and directly avoided in apprehension of both the undertaking and the results by others. This publication shows the evaluation in its complexity and strives to dismantle this unflattering picture of policy evaluation.

Evaluation comprises of many steps and can bring numerous benefits to all stakeholders, starting with the obvious empirical evidence and assessment of the given policy, continuing through creation of new relationships among the stakeholders, looking for the common ground in terms of policy meaning and implementation, and even increasing engagement of the recipients of the policy. Evaluation can be a tool of public debate, a tool contributing to further development of policies, and a sign of accountability.

Evaluator is not a major critic of all that is in place; evaluator is a critical friend who comes to point out all the positive as well as all the negative aspects of any policy, keeping in mind that the further development and improvement of any policy is the main objective, and ideally also providing recommendations to achieve that.

Youth policy field is a complex one. It often requires cross-sectoral cooperation, differentiated funding schemes, and very specific approach when implementing concrete measures. All of these are reasons for evaluation to not only take place, but to be planned for along with the policies to be implemented, to be conducted continuously as a way of monitoring the progress and adjusting any of the many components along the way if necessary.

Evaluation process does not have to be an enormous undertaking suitable only for the largest stakeholders in the field. It can be planned as an internal process ensuring the quality control and development within a single NGO, youth club, or a leisure time centre. Evaluation findings do not have to be shocking and demand the change of the whole system, they can point out small changes to bring about a better policy implementation or performance in the future.

Hopefully, this publication contributes to spreading the idea of evaluation across the youth sector, inviting various actors to start as many evidence-based processes on the local, regional, national and international levels, as possible. In times of fake news and disinformation, evidence-based decision-making is more important than ever, and so is the evidence-based and informed public debate. Evaluation is a tool to support these goals and help develop youth policies further.

# Sources

**Bárta, O., Fennes, H., Gadinger, S., Böhler, J., Mayerl, M. 2019.**

RAY-LTE: Long-term Effects of Erasmus+ Youth in Action on Participation and Citizenship.  
Online, available at <https://www.researchyouth.eu/download/?id=499>

**Backett-Milburn, K. & Wilson, S. 2000.**

Understanding peer education: insights from a process evaluation.  
In Health Education Research, Volume 15, Issue 1, Pages 85–96,  
<https://doi.org/10.1093/her/15.1.85>

**Bammer, D., Fennes, H., Karsten, A., Gadinger, S., Mayerl, M., Böhler, J. 2017.**

Exploring Erasmus+: Youth in Action. Effects and Outcomes of the Erasmus+: Youth in Action Programme from the Perspective of Project Participants and Project Leaders. Transnational Analysis 2015/16 – Main Findings. Vienna: GENESIS.  
Online, available at <https://www.researchyouth.eu/download/?id=198>

**Böhler, J., Fennes, H., Karsten, A., Mayerl, M., Pitschman, A., Roth, C.A. 2019.**

Exploring Erasmus+: Youth in Action. Effects and Outcomes of the Erasmus+: Youth in Action Programme from the Perspective of Project Participants and Project Leaders. Transnational Analysis 2017/18 – Main Findings. Vienna: GENESIS.  
Online, available at <https://www.researchyouth.eu/download/?id=497>

**Commission of the European Communities. 2004.**

Impact Assessment Integrating Ex-Ante Evaluation of the “Youth in Action” Programme.  
Brussels: Commission of the European Communities.  
Online, available at [https://ec.europa.eu/youth/sites/youth/files/youth-in-action-ex-ante-impact-2004\\_en.pdf](https://ec.europa.eu/youth/sites/youth/files/youth-in-action-ex-ante-impact-2004_en.pdf)

**Commission of the European Communities. 2008.**

Final evaluation of the “Youth” Community action programme (2000-2006) and of the Community action programme to promote bodies active at European level in the field of youth (2004-2006).  
Brussels: Commission of the European Communities.  
Online, available at <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM%3A2008%3A0398%3AFIN%3AEN%3APDF>

**Council of the European Union. 2018.**

Resolution of the Council of the European Union and the Representatives of the Governments of the Member States meeting within the Council on a framework for European cooperation in the youth field: The European Union Youth Strategy 2019-2027.  
Official Journal of the European Union.  
Online, available at <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:C:2018:456:FULL>

**Charles University. 2018.**

Internal Evaluation Board.  
Online, available at <https://cuni.cz/UKEN-752.html>

**Chen, Huey T. 2015.**

Practical Program Evaluation. Theory-Driven Evaluation and the Integrated Evaluation perspective.  
Thousand Oaks: SAGE Publications.

**Chiwara, Richard M. & Hussein Y. Ali. 2018.**

END OF PROJECT EVALUATION - JOINT PROGRAMME ON YOUTH EMPLOYMENT.  
UNDP Evaluation Resource Centre.  
Online, available at <https://erc.undp.org/evaluation/evaluations/detail/9365>

**De Vaus, David. 2014.**

Surveys in Social Research.  
London: Routledge, Taylor & Francis Group.

**De Vaus, David. 2002.**

Analyzing Social Science Data.  
London: SAGE Publications.

**Directorate-General for Education, Youth, Sport and Culture. 2017.**

Study on the impact of transnational volunteering through the European voluntary service.  
Luxembourg: Publications Office of the European Union.  
Online, available at <https://publications.europa.eu/en/publication-detail/-/publication/01a810b3-3712-11e7-a08e-01aa75ed71a1>

**Directorate-General for Education, Youth, Sport and Culture & ICF. 2017.**

Combined evaluation of Erasmus+ and predecessor programmes. Final report: main evaluation report (Volume 1).  
Luxembourg: Publications Office of the European Union.  
Online, available at <https://publications.europa.eu/en/publication-detail/-/publication/3d783015-228d-11e8-ac73-01aa75ed71a1/language-en>

**Education Review Office. 2016.**

Effective Internal Evaluation for Improvement.  
Online, available at <https://www.ero.govt.nz/assets/Uploads/Effective-internal-evaluation-for-improvement.pdf>

**Eesti Noorsootöö Keskus. 2019.**

Youth Monitor.  
Online, available at <https://www.noorteseire.ee/en/indicators>

**ESS ERIC. 2019.**

European Social Survey.  
Online, available at <https://www.europeansocialsurvey.org/>

**European Commission. 2018a.**

REPORT FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS. Mid-term evaluation of the Erasmus+ programme (2014-2020).  
Brussels: European Commission.  
Online, available at <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0050&from=FR>

- European Commission. 2018b.**  
COMMISSION STAFF WORKING DOCUMENT. Mid-term evaluation of the Erasmus+ programme (2014-2020). Brussels: European Commission.  
Online, available at <https://ec.europa.eu/assets/eac/erasmus-plus/eval/swd-e-plus-mte.pdf>
- European Commission. 2014.**  
Erasmus Impact Study confirms EU student exchange scheme boosts employability and job mobility. Press release.  
Online, available at [https://europa.eu/rapid/press-release\\_IP-14-1025\\_en.htm](https://europa.eu/rapid/press-release_IP-14-1025_en.htm)
- Fennes, H., Hagleitner, W., Helling, K., Rosenthal, A., Chisholm, L. 2011.**  
Research-based Analysis of Youth in Action. Results from the first series of surveys with project participants and project leaders between October 2009 and June 2010. Transnational Analysis. Innsbruck.  
Online, available at <https://www.researchyouth.eu/download/?id=28>
- Fennes, H., Gadinger, S., Hagleitner, W., Chisholm, L. 2012.**  
Exploring Youth in Action. Effects and outcomes of the Youth in Action Programme from the perspective of project participants and project leaders. Transnational Analysis 2011 – Executive Summary. Innsbruck.  
Online, available at <https://www.researchyouth.eu/download/?id=31>
- Fennes, H., Chisholm, L. 2013.**  
Youth in Action makes a difference!  
Vienna: Interkulturelles Zentrum.  
Online, available at <https://www.researchyouth.eu/download/?id=33>
- Frechtling, Joy, Melvin M. Mark, Debra J. Rog, Veronica Thomas, Henry Frierson, Stafford Hood, Gerunda Hughes. 2010.**  
The 2010 User-Friendly Handbook for Project Evaluation. Rockville: Directorate for Education and Human Resources.
- Goffman, E. 1956.**  
The Presentation of Self in Everyday Life. University of Edinburgh.
- Hora, Ondřej, Miroslav Suchanec, Martin Žižlavský. 2014.**  
Evaluacní výzkum. Brno: Masarykova univerzita.
- Independent Evaluations Office. 2019.**  
Independent Evaluations Office.  
Online, available at <http://web.undp.org/evaluation/evaluation-office.shtml>
- Kloosterman, P., Giebel, K., Senyuva, O. 2007.**  
T-Kit no.10. Educational Evaluation in Youth Work. Strasbourg: Council of Europe.  
Online, available at [https://pjp-eu.coe.int/documents/42128013/47261233/T-Kit\\_10.pdf/8d85c6ac-05e5-4715-8f43-0f9c3018772a](https://pjp-eu.coe.int/documents/42128013/47261233/T-Kit_10.pdf/8d85c6ac-05e5-4715-8f43-0f9c3018772a)
- Knox, Alan B. 2002.**  
Evaluation for Continuing Education. San Francisco: Jossey-Bass.
- McCoshan, A., Koppert, S., de Bruin, G., Warmerdam, S. McKenna, K., Day, L. 2011.**  
Youth in Action Interim Evaluation. Final Report. Rotterdam: ECORYS.  
Online, available at [https://ec.europa.eu/youth/sites/youth/files/youth-in-action-interim-evaluation-2011\\_en.pdf](https://ec.europa.eu/youth/sites/youth/files/youth-in-action-interim-evaluation-2011_en.pdf)
- Miller, R.L. & Campbell, R. 2006.**  
Taking Stock of Empowerment Evaluation: An Empirical Review. In American Journal of Evaluation, Vol. 27, No. 3, pages 296-319.  
Online, available at <https://pdfs.semanticscholar.org/236a/2854a6b1a611f835efc2b6fb59d9e773d2c1.pdf>
- Newell, S. & Graham, A. 2012.**  
Using an empowerment evaluation approach with community-based programs. Reflections from the front line. In Evaluation Journal of Australasia, Vol. 12, No. 2, pages 15-27.  
Online, available at <https://www.aes.asn.au/images/stories/files/Publications/Vol12No2/3.pdf>
- OECD Governance. 2019.**  
Seven Key Findings from the Youth Governance Survey. Online, available at <https://www.slideshare.net/OECD-GOV/seven-key-findings-from-the-youth-governance-survey-152198866>
- Patton, Michael Q. 2002.**  
Qualitative Research and Evaluations Methods. Thousand Oaks: SAGE Publications.
- Rossi, Petr H., Marke W. Lipsey, Howard E. Freeman. 2004.**  
Evaluation: A Systematic Approach. Thousand Oaks: SAGE Publications.
- Shulha, L. M., Whitmore, E., Cousins, J. B., Gilbert, N., Al Hudib, H. 2015.**  
Evidence based principles to guide collaborative approaches to evaluation: Technical report. Ottawa: Centre for Research on Educational and Community Services, University of Ottawa.  
Online, available at [https://crecs.uottawa.ca/sites/crecs.uottawa.ca/files/shulha\\_et\\_al\\_2015.pdf](https://crecs.uottawa.ca/sites/crecs.uottawa.ca/files/shulha_et_al_2015.pdf)
- Sida's Evaluation Group. 2018.**  
Sida's Evaluation Handbook. Stockholm: Sida's Evaluation Group.
- Smutek, Martin. 2005.**  
Evaluace sociálních programů. Hradec Králové: Gaudeamus.
- Tilburg University. 2019.**  
European Values Study.  
Online, available at <https://europeanvaluesstudy.eu/>
- Thonon, C.B. & López, M. Á. G. 2014.**  
Youth Peace Ambassadors. Human dignity and peace-building by young people in Europe. External Evaluation Report of the Youth Peace Ambassadors Project.  
Online, available at <https://rm.coe.int/1680460485>
- UNESCO. 2014.**  
8th UNESCO Youth Forum Evaluation.  
Online, available at <https://unesdoc.unesco.org/ark:/48223/pf0000229162>
- United Nations Development Programme. 2019.**  
UNDP Evaluation Guidelines. New York: UNDP.
- Weiss, Carol H. 1998.**  
Evaluation. New Jersey: Prentice-Hall.
- Wholey, Joseph S., Harry P. Hatry, Kathryn E. Newcomer (eds.). 2010.**  
Handbook of Practical Program Evaluation. San Francisco: Jossey-Bass.
- YOUTHCONF. 2019.**  
Youth Goals.  
Online, available at <http://www.youthgoals.eu/>

