

Introduction

The INTOSAI Working Group on the Evaluation of Public Programs and Policies was established in 1992. Its objective is to assist SAIs that conduct evaluations, or wish to do so, by providing them with reference frameworks, methodological tools and practical recommendations to implement this specific approach, which is different from other forms of control and audit. For this reason, the working group was placed under the authority of Intosai's Objective 3 "knowledge sharing" (KSC).

It currently includes 22 SAIs. Since its creation, it has been chaired by the French Court of Auditors. As performance audit is the form of audit closest to evaluation, Brazil has obtained observer status as Chair of the Subcommittee on Performance Audit (PAS) of the Professional Standards Committee.

In 2010, after 18 years of work, the group finalized a first document on evaluation. It was validated by the KSC and then definitively approved at the 20th INCOSAI Congress in Johannesburg. This document consisted of a broad survey of program evaluation in the various SAIs and aimed, through a "bottom-up approach", to present a general definition of evaluation and to propose general recommendations for evaluation planning. But it was generally limited to an inventory of the different SAIs' practices in this area. This first document led the group to propose to deepen the definition and methodology of the evaluation. It then sent a new questionnaire to the various SAIs and received 44 responses from SAIs interested in program evaluation. In 2012, at the 4th KSC meeting in Luxembourg, the group decided to prepare a methodological guide to complete the first volume by refining the evaluation concept and offering a wide range of examples, advice and recommendations on evaluations. It also decided to establish links between the EWG and the PAS (Sub-Committee on Performance Audit) since the distinction between these two approaches is sometimes difficult to understand for some SAIs and program evaluation is sometimes perceived as a component of performance audit.

This methodological guide is the result of two years of work and discussions. Between April 2013 and January 2015, it was amended several times and discussed among EWG members. In addition, the KSC Annual Steering Committees and INTOSAI Steering Committees (in 2013 and 2014) have received numerous reports on it. The guide aims to define precisely the main characteristics of an evaluation (estimation of the value of a program or policy, involvement of stakeholders, etc.) and what makes it different from a performance audit. It also describes the evaluation methodology that SAIs should adopt (criteria for selecting the policy to be evaluated, outcomes to be measured, date of commencement of the evaluation, different steps in the evaluation process, etc.) and how to proceed with the different stakeholders in a public policy. Finally, it focuses on evaluation publication issues and the boundaries between program evaluation and policy intervention, a barrier that an evaluation cannot overcome under any circumstances.

However, the purpose of this methodological guide is not to define a standard for a new audit category, as there are different evaluation practices among SAIs beyond the common principles. Moreover, it would be unnecessary and counterproductive to distinguish "pure and perfect evaluation" from other existing practices. Instead, the guide should encourage individual SAIs and the audit community to take an interest in program evaluation and provide them with guidelines to help them carry out evaluations in a scientific and independent manner. The guidelines in this document are therefore intended to set out the principles to be applied, the pitfalls to avoid and describe good practices rather than to establish strict rules. The objective is to provide a solid basis for improving the capacities of SAIs and auditors in this area. Although a guide is not enough to strengthen the organization and management of evaluations, it is the first step in the long process of creating a rigorous evaluation methodology in SAIs. It will enable SAIs to address methodological difficulties and help them to find answers in a very extensive literature. This guide is a methodological and practical tool.

With regard to the link between this guide and the standards applicable to performance auditing, Part 1 of ISSAI 3000 already refers to program evaluation by attempting to define precisely the performance audit (called "performance audit" in ISSAI 3000) and its relationship to other audit categories:

- The performance audit examines the economy (are public and private resources used in accordance with good administrative and financial management practices?), efficiency (are the resources used in the best way?), effectiveness of a program (are the objectives achieved with the resources used and can the results be observed?). It focuses on measuring the performance of a program or administrative unit by examining the relationship between resources and outputs on the one hand, and outputs and outcomes on the other. It may also question the objectives of a program and ask for alternative policies. However, this is not an obligation and does not fall within the scope of its main objective. Performance measurement, normally refers to the examination of predefined objectives. It is not intended to question these objectives and, a fortiori, to request that a program or policy be modified or suspended.
- Program evaluations are ad hoc and systematic studies designed to assess a theme in a comprehensive way, using more or less complex scientific approaches. The main objective is not to assess the effectiveness of a program or policy, its efficiency or, more generally, the causal relationship between the resources used and the intended results, even if these elements are part of the basis of a program evaluation. This type of evaluation tends, on the one hand, to analyze the long-term socio-economic impact of a program or policy and, on the other hand, to question its value. In addition, program evaluations do not target restricted programs, but always a broad set of programs with a common goal, or a policy implemented several years earlier.

Taking up the different categories of evaluation recalled in ISSAI 3000, while monitoring the economy, efficiency or effectiveness of a program or policy is at the heart of performance auditing, the key to program evaluation is to assess the overall and long-term socio-economic impact of a policy (which often involves taking into account other policies having effects in the same field) but also its overall relevance.

However, both performance audits and program evaluations aim to go beyond regularity control, as they are not interested in the implementation of legal provisions and do not judge the legality of the means used. Rather, they are dedicated to measuring results against means.

For some, based on the ISSAI 3000, performance audit appears to be a very broad evaluation comprising different categories of controls, including program evaluation. It appears that program (or policy) evaluation is a very particular type of performance audit with characteristics and objectives that require further explanation.

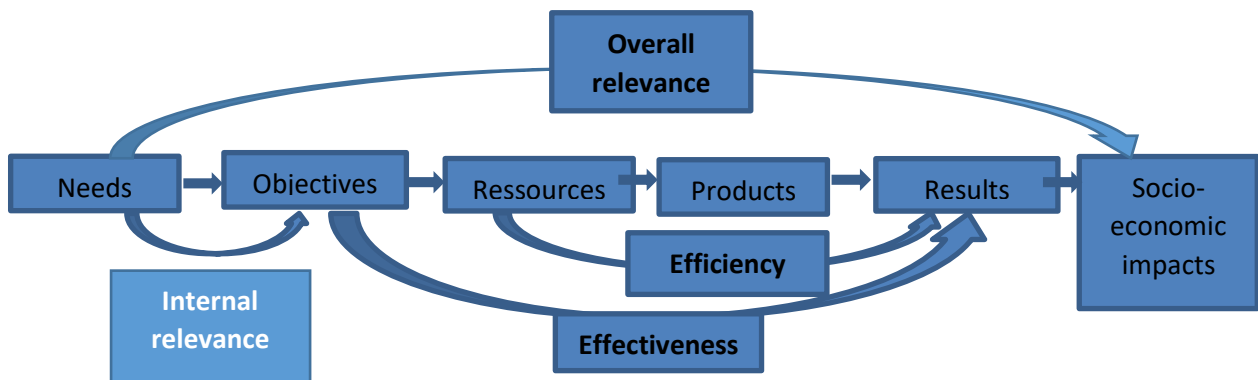
This is what motivated the development of this methodological guide. As evaluations become increasingly important for SAIs, they need common guidelines to encourage their role and strengthen their independence in this area.

I. Definition and objectives of an evaluation

An evaluation is the examination of a program or public policy to assess the value of that program or policy. An evaluation analyses as systematically as possible its objectives, implementation, achievements and socio-economic impacts and measures its performance and objectives in order to

finally assess its relevance. Evaluation is therefore increasingly important for public debate, as public opinion and political leaders must make decisions based on evidence.

As both program evaluation and performance audit are intended to measure performance, it is not easy to differentiate between the two procedures. However, the following diagram provides information on their key elements:



Some characteristics are common to program evaluation and performance audit, namely performance measurement. These two surveys examine the causal relationships between public action and its effects (use of a socio-economic impact model distinguishing between objectives, implementation, outputs, results and socio-economic impacts), whether these effects are direct or indirect. Therefore, they have the same objective: to measure and categorize the effects of a program. But even from this point of view, there is a slight difference between these two approaches:

- As ISSAI 3000 points out, "performance auditing assesses economy, efficiency and effectiveness, and includes:
 - review of savings in administrative activities, in accordance with sound administrative principles and practices,
 - verification of the efficiency of resource use,
 - verification of the effectiveness of the results in relation to the objectives pursued by the audited unit. It is a question of determining whether the defined objectives are being achieved and to what extent they are being achieved through the means employed. In addition, measuring a program's performance must also take into account its unintended impacts and effects.

Since performance auditing focuses mainly on the causal link between, on the one hand, the instruments used and the resources employed and, on the other hand, the direct and indirect effects, "performance auditing is based on decisions taken or objectives set by the legislature". The main objective of the performance audit is to report on "the achievements of a program, including progress towards predefined objectives".

- On the contrary, program evaluation is more concerned with:
 - the internal relevance of a program or policy, i.e. the value of its objectives in meeting targeted social, economic or environmental needs;
 - the overall relevance of a program or policy. The question is whether the program or policy was useful, taking into account all the direct (outcomes) and indirect, or even

unexpected, effects of the impacts, and the needs that the program was intended to address.

In short, program evaluation is not limited to predefined objectives and makes it possible to question the objectives set out in the legislation. The main difficulty is to question the value of a program or policy, not the effectiveness of its instruments.

However, the two approaches are complementary and together constitute two decisive components for judging a program or policy. Measuring the effects of a policy makes it possible to estimate the efficiency and effectiveness of the policy being evaluated. These elements are essential components of performance audit. The latter will then be used to contribute, among other elements (taking into account other policies in place in the same scope or dealing with the same issues, examining alternative policies, etc.), to assessing the overall relevance of a policy. This approach is characteristic of public policy evaluation. However, this does not mean that a performance audit cannot focus on the overall relevance of a program at the end of its investigation or assess the value of a policy. However, these are optional elements that are not part of the main objectives of a performance audit, as explained in ISSAI 3000, which is the result of a review of different practices. The majority of performance audits carried out in SAIs focus on organization, processes and results, not on socio-economic impacts and policy objectives. Of course, there are exceptions. In Germany, for example, the evaluation of project results is an essential element of performance evaluation. Conversely, in Switzerland or France, the evaluation is completely separate from the performance audit and is carried out using a specific method.

In the context of a "performance audit", the main characteristic of a program evaluation is therefore the existence of a value judgment on the objectives and socio-economic impact of a policy, as well as the desire to improve the overall relevance of this policy through evaluation. The objective of an evaluation is therefore to inform the public about public actions and their effects, and to assess the value of these actions in order to help decision-makers improve their overall relevance and socio-economic impacts.

It is questionable whether this assessment of the value of a program or policy does not go beyond the mandate of SAIs and may lead them to enter the political debate. In principle, this is neither possible nor desirable. So far, in all countries where SAIs have created an independent evaluation, the answer is about the same. SAIs propose an independent study on the objectives or socio-economic impacts of a program or policy. In doing so, on the one hand, they allow the public to form its own opinion on public actions, and on the other hand, they help political decision-makers to pursue or modify a policy based on concrete elements (and not on mere intuitions or opinion polls). This contribution to the democratic debate is neutral and fact-based. It proposes a reflection on a given policy through objective analyses and alternatives, but without advocating a single solution or opting for a political party. Consequently, SAIs do not interfere in the political debate that divides political parties into actions to be taken, nor do they force political authorities to take a decision on a program or policy.

For example, in Germany, the SAI highlights the audit findings regarding underlying assumptions or the effects of policy measures. And in recent years, it has developed a systematic evaluation of legislative acts in order to inform decision-makers about the incentive effects of legislation. Nevertheless, it refrains from interfering in political decisions. Even in Switzerland, where the tradition of evaluation is particularly developed, recommendations are not a mandatory element of evaluations. France offers a concrete implementation of this distinction with a recent assessment of an environmental policy (on energy and its effects on the climate). This evaluation concludes with proposals on the importance given to renewable energies in this policy, but does not at all call into question the need for this

environmental policy as such. For all SAIs interviewed on their evaluation practice, it is appropriate for political authorities to draw conclusions from such an evaluation, which can lead to:

- the continuation of the policy or its expansion / acceleration;
- the modification, correction or termination of the policy.

From this main characteristic (the assessment of the overall relevance of a policy) flow other characteristics specific to program evaluation:

- The program or policy studied cannot be made up of a single administrative unit or measure. A program or policy is a range of measures, human, legal and financial resources, or even organizations working towards a common objective, that have been in place for several years. For example, the whole environmental policy cannot be assessed, nor can a tax incentive to encourage the purchase of cleaner cars. Only an environmental policy such as action to promote energies that are less harmful to the climate can be assessed. Stakeholders are automatically included in the evaluation (support groups, interviews, etc.) as the program has a socio-economic impact on them and their views must be taken into account to obtain a comprehensive evaluation. It is also important to take into account the role of other authorities (international, national, regional and local) and civil society actors to avoid focusing only on a given administrative unit.
- Different scientific research methods are used (qualitative studies and social data, quantitative indicators) since the objective of an evaluation is to encompass all socio-economic impacts of a program (including behavioral and subjective impacts).
- A report and a transparent description of the method used and the results produced must be published.

In conclusion, program evaluation and performance audit share the same objective: to go beyond regularity control and measure in one way or another the effects of a program or policy. But while performance measurement is limited to a simple verification of results, which can be compared to established objectives and whose level of success is measurable (especially in terms of costs), evaluation in the strict sense of the term goes further to assess the overall relevance and conduct of the program or policy evaluated. Therefore, this is a very particular type of performance audit.

This value assessment can add a new dimension to the work of SAIs, alongside their other missions. This is what the public (elected representatives, decision-makers, managers, media, opinion, etc.) is entitled to expect from the evaluation of public policy conducted by an SAI. Based on the different aspects studied, it should be able to reach a conclusion on the overall relevance, or lack thereof, of a given policy and help policy makers to question the existence of alternatives to that policy.

II. Evaluation actors and institutional environment

SAIs are not the only actors that can conduct evaluations:

- Evaluations can also be carried out by independent academic organizations and private consulting firms, on their own initiative or at the request of a policy maker such as Parliament.
- Administrative institutions, such as control and audit bodies, may carry out evaluations at the request of their government authorities.
- Finally, the actors responsible for implementing a public policy may also decide to evaluate it or manage their evaluation.

- In addition, an SAI conducting an evaluation in an area that has already been studied should collaborate with previous evaluators and take into account their work.

Nevertheless, SAIs are the natural actors in public policy evaluation, as they have the necessary independence and have acquired, during their other missions, the methodology and essential knowledge of public policy. Unlike administrative institutions or private bodies, they do not have to prove their objectivity and can guarantee their independence from government or private interests.

An SAI may decide to conduct an evaluation on its own initiative (voluntary evaluation) or at the request of a public authority, Parliament, or the executive (called the "sponsor"). In this case, the SAI may engage in a dialogue with the sponsor to determine the precise outlines of the policy to be evaluated and the evaluation issue. And although an SAI takes into account the views of stakeholders when planning its audits, it nevertheless decides on the scope of the audit and its process, manages the conduct of the evaluation itself and prepares its conclusions. The SAI should therefore refuse to respond to a request from a public authority or to initiate a program evaluation if it fears that its independence may be threatened.

In all cases, SAIs should consider other possible sources of assessment for the subject concerned.

III. Evaluation methodology and planning

A. Choice of the subject of the evaluation

The evaluation may focus on:

- public policy (which may be composed of several programs);
- a specific public action program including legal and financial provisions, the actions of different organizations, human or financial resources aimed mainly at achieving specific objectives.

Several criteria can be used to select a program or policy to be evaluated, including the importance of the policy, the ability to measure its impact and the time elapsed since its launch.

In order to decide whether it is possible to evaluate a particular program or policy, the SAI carries out a feasibility study. The objective is to define the framework of the evaluation and the conditions under which it can be implemented.

First criterion: Importance of the program or policy to be evaluated

Two pitfalls should be avoided when choosing the subject of an evaluation.

The first trap would be to choose a policy or program that is too broad (environmental policy, employment policy, education policy). The scope of the policy to be evaluated should not be too broad. Despite possible demands and temptations, it is not really possible to evaluate an education policy, a housing policy, or any other overall policy as such when the evaluation must meet the requirements defined above. It is even more difficult to assess the state of a sector as a whole: the state of schools, social protection, defense or justice, for example.

The second trap would be to target a very specific project or public action (infrastructure, response mechanism, tax mechanism, etc.). In fact, it is not desirable that the program or policy to be evaluated be too narrow.

- The socio-economic impact of the policy on society may be too limited and an SAI that conducts only a few public policy evaluations each year cannot disperse its teams.
- The objectives and effects of highly targeted policies are limited and may be overestimated if they are not integrated into a more meaningful and broader set. It would be more interesting to use resources to compare the effects of related provisions according to their objectives and independent provisions with the same objective.
- Due to a lack of sufficient comparability, it is difficult to make relevant parallels with situations abroad.
- Above all, such a review would be closer to performance audit than evaluation in the strict sense of the term.

The "importance" of a policy can therefore be determined by the following criteria:

- its budget, i.e. the amount of public funding allocated to this public policy;
- the number or importance of stakeholders, or the complexity of their role;
- the scope of its expected effects on beneficiaries and society;
- the complexity of a policy due to the multiplicity of stakeholders and the difficulty of assessing its effects;
- the symbolic importance of politics to public opinion. For example, a renewable energy policy (and not energy policy as a whole) or an alcohol policy (and not health policy) can be evaluated.

Second criterion: possibility of measuring the majority of the "effects" of the policy

The idea of measurable socio-economic effect/impact is essential to the evaluation. On the one hand, examining the overall relevance of a policy requires measuring its effects by linking them to other aspects such as the costs and organization of the policy: effects can be assessed in both quantitative and qualitative terms (i.e. a value judgment can be made). On the other hand, this measure is often technically complicated and requires the longest work.

The second criterion to be taken into account is therefore the existence of measurable effects that can be attributed to the policy implemented.

Two types of effects can be distinguished:

- Direct and immediate or short-term "effects", which have an impact on direct beneficiaries, are "outcomes".
- Delayed or medium- to long-term "effects", often felt by people other than direct beneficiaries, are "socioeconomic impacts".

The distinction between results and socio-economic impacts is essential. It is therefore strongly recommended to apply this distinction when evaluating. The measurement of long-term effects is characteristic of program evaluation and goes further than the measurement made during performance audits.

However, the purpose of public policy evaluation is not only to measure the effects associated with pre-defined objectives, but also to evaluate all the effects of a public policy, including unintended effects. A second distinction is therefore necessary to differentiate between the two types of effects to be assessed:

- Effects related to the objectives defined by the policy (the "expected effects"). The evaluation is based primarily on the measurement of predicted effects and assumes that the objectives for these effects are stated and can be analyzed.
- Other effects, even if they were not anticipated by the stated objectives: "unforeseen effects" that were not explicitly expected and that are "positive" or "negative" (sometimes called "perverse effects"). The evaluation must take into account all the long-term and unexpected effects of a policy in order to be comprehensive and not be limited by its objectives.

Finally, evaluation should not only consider the objective effects of a policy, but also the effects "felt" by policy actors and beneficiaries. Measuring the objective and perceived effects requires placing public policy "stakeholders", i.e. the administrations and public bodies concerned, potential beneficiaries and third parties involved (organizations and individuals) at the heart of the evaluation process, as it is their positions and opinions that must be described and evaluated.

The objective of the evaluation is to measure all effects and describe the causal links between the program or policy evaluated and these effects. It is essential that the evaluation aims to attribute the observed effects to the instruments of the policy implemented, otherwise it is impossible to confirm the overall relevance of a policy. It must therefore be based on a causal reasoning that can prove the link between political action and a measured effect. For some, evaluation must "model the policy", i.e. identify (or reconstruct) a system of interaction between the different measures taken as part of a policy and their measured consequences, i.e. their effects.

The table below provides a systematic approach for all effects under evaluation. It will be difficult to complete satisfactorily, but it must, in the background, be one of the ambitions of the evaluation.

	Direct effect ("result")	Indirect effect ("socio-economic impact")
	Objective / Feeling	Objective / Feeling
Expected effect		
Unintended Effect		

Third criterion: time elapsed since the policy was launched

From a theoretical point of view, an evaluation can be launched at three different times.

- An ex ante evaluation is carried out before the launch of a project or program, or before the creation of an organization, which happens very rarely in the case of SAIs.
- A concomitant evaluation shall be carried out during the implementation and execution of the action to be evaluated. The decision to conduct this type of evaluation can be made at the time of the launch of the program or legislation formalizing public policy.
- An ex-post evaluation is a retrospective analysis, a revision. This is the most common for SAIs. Sometimes, this ex-post evaluation is explicitly provided for as soon as a program is launched.

Although there are three types of evaluation, SAIs are most often mandated for "ex-post" evaluations, i.e. a few years after the launch of the program or policy.

Before implementing an evaluation, it is necessary to wait two or three years from the start of a program in order to have sufficient data and avoid provisional results. On the other hand, since indirect and long-term effects are essential components of evaluation, it is also crucial to evaluate them. The policy chosen for the evaluation must have been in place for less than seven or eight years. However,

from this point of view, it is preferable to evaluate an older policy than a fairly recent policy, recently adopted or implemented, or whose characteristics have recently changed, because its effects would be difficult to define and measure.

By allowing a little time to pass, it is also easier to define a "state of the world without this policy" (also called "counterfactual experience" or "state of the world before a policy is launched") and to compare it with the current state, in order to deduce the effects of the policy implemented.

B. Project construction with stakeholders

Evaluation is a particular process in particular because it is based on a close relationship, or even a co-construction of the approach with the stakeholders involved in the policy or program (whether or not they fall within the SAI's supervisory competence) or direct or indirect beneficiaries of the actions evaluated, or simply affected by the policy. The definition of stakeholders is also particular in that the objective is not to correct administrative malfunctions or even to sanction individual failures, but rather to improve a policy.

- The actors of the policy evaluated are the public authorities that have decided and implemented it. Administrative authorities should be considered as a separate case. They are responsible for the policy and are always considered as stakeholders, whether at local or national level. As these authorities are usually part of the SAI's supervisory competence and are familiar with its other interventions, it is important that the evaluation team develop a new relationship with them, different from that which exists in the context of regular audits and based on mutual trust and co-construction (e. g. working together to achieve a common objective, providing answers to evaluation-related questions). It should be stressed that the estimation of the value provided on the overall relevance of the policy is addressed more to the political authorities that have requested it than to the administrative authorities that have implemented it, and that they should therefore not feel that they have been put in an uncomfortable position by the objectives of the evaluation. To gain and maintain this trust, it is also important to ensure that the SAI does not focus solely on the discovery of administrative malfunctions and secondary irregularities as it would during an audit: this is not at the heart of what the evaluation seeks to achieve.
- Local elected representatives who could be considered as stakeholders when the policy evaluated involves a contribution from regional or local autonomous bodies.
- Legislators can sometimes be considered as stakeholders if, for example, they have voted for legislative provisions that form the basis or framework of the public policy being evaluated, let alone when these provisions have established a mandatory mechanism for evaluating the actions undertaken. But this issue is controversial, as Parliament is a political authority, but also the recipient of the evaluation: this dual role could jeopardize the independence of the evaluation process. In addition, it would lead to the almost systematic integration of Parliament among stakeholders, which would be difficult to understand (a stakeholder is concerned with the policy implemented and not with all policies in general) and to achieve (how would Parliament's representatives be chosen?) For this reason, some SAIs, in Switzerland for example, never consider Parliament as a stakeholder. Nevertheless, if the SAI acts on a request that does not come from Parliament, Parliament may be considered a stakeholder because of its membership of the supervisory body, whereas in the case of a voluntary evaluation at the SAI's initiative, it would be more difficult for the SAI to integrate parliamentary representatives among its stakeholders.

Other stakeholders are private organizations (NGOs, companies, professional bodies, trade unions, etc.) interested in the implementation of the policy and representing the beneficiaries of the policy being evaluated. The development of the list of stakeholders is a sensitive point in the evaluation process. These stakeholders rarely fall within the SAI's supervisory competence, are generally not accustomed to working with the SAI and are under no obligation to participate in its activities. It is therefore important to ensure that these stakeholders understand that it is in their interest to be constructively involved in the evaluation, and to create an atmosphere of trust when interacting with them. It is also advisable to correct any incorrect perception they may have of the SAI as a supervisory body in the event that the evaluation is carried out in a participatory approach.

Two pitfalls should be avoided with regard to stakeholders: the list of stakeholders should not contain any major omissions (inappropriate exclusions could undermine the scope of the evaluation findings, the chances of their implementation and, above all, lead to a partial vision of the policy), but neither should it be too long (which could make the process too complex and unmanageable).

Unlike most performance audits, stakeholder involvement is a key factor in the evaluation, as stakeholder representatives support the SAI in the evaluation process. They must adopt a constructive attitude towards the evaluation process, its objective being not to reveal and correct dysfunctions, but to help decision-makers to meet social and political needs. Stakeholders must therefore

- Be involved in the choice of the precise subject of the evaluation, the timetable and the methodology
- Actively participate in the evaluation
- Be kept informed during or at the end of the process
- Have a role to play in the post-evaluation decision-making process

It should nevertheless be clear that, although an interactive dialogue between the sponsor and the SAI carrying out the evaluation is fundamental to the evaluation process, this communication does not compromise the independence and impartiality of the SAI. In an evaluation, it is important that the SAI remains independent of political influence and does not have to respond to the wishes of particular political interests. This point should be clarified at the beginning of the preparation of the evaluation.

All countries with a long tradition of evaluation, such as Switzerland, stress the importance of letting the SAI have the final say in the evaluation process.

C. Evaluation planning

Once the subject of the evaluation has been pre-identified, evaluation planning can begin. It is divided into several stages:

- Feasibility study
- Evaluation Management
- Finalization of the results

These steps are not specific to the evaluation program per se, they are the same as for all performance audits. But there are some particular problems inherent in the evaluation process that should be highlighted.

C.1. feasibility study

Before deciding whether to carry out a voluntary evaluation or to accept a request for an evaluation from an external authority (sponsor), the SAI should consider the feasibility of its own approach, or the

approach it is being asked to adopt. The feasibility study addresses different elements and addresses all issues that need to be addressed to ensure the success of the evaluation to the extent possible.

- definition of the subject and scope of the evaluation (see III-A)
- definition of stakeholders (see III-B)
- definition of the issues to be addressed in the evaluation
- organization of the evaluation process and planned timetable for the evaluation
- choice of methodology and scientific instruments
- identification of human and financial resources.

The different responses provided are examined in a feasibility study, the results of which are presented to the SAI's decision-makers. Based on the results of this study, the SAI decides whether or not to continue the evaluation process and launch the evaluation phase itself.

As mentioned in ISSAI 3000, the definition of the audit objectives, scope and methodology during a preparatory study is not specific to program evaluation. But the feasibility study is particular during the evaluation process in that the objective is to define whether or not a program can be evaluated - according to the above-mentioned criteria - and whether or not it is possible to organize an independent evaluation with stakeholders.

The feasibility study phase is relatively straightforward when the SAI conducts a voluntary evaluation: In this case, the SAI does not need an interactive dialogue with stakeholders (although it is not systematically to be excluded) and can itself use quantitative and qualitative data as well as external data to carry out the evaluation, allowing it to decide on the above-mentioned evaluation criteria.

When the evaluation project responds to an external request, the feasibility study is more complex, as it must take into consideration not only the information already mentioned, but also the intentions of the sponsors in terms of problems to be solved, questions to be asked, timetable to be respected and expected results. Nevertheless, the SAI team responsible for the feasibility analysis should in this case not only take into account the intentions of the sponsors, but also inform them of the methodological, organizational and time requirements of the evaluation. The interactive dialogue between the sponsor and the SAI carrying out the evaluation is essential to formulate the key evaluation questions and plan the approach to be taken in an appropriate way. It may also lead to the conclusion that an evaluation on the subject proposed by the sponsor is not possible (because of its very nature or the conditions requested by the sponsor). In all cases, the SAI is solely responsible for the decision to carry out an assessment and will have to respond in the negative to the request when the criteria relating to the subject and the conditions necessary for the process are not met.

A final point concerns the number of evaluations carried out in response to external requests, which should not jeopardize the balance between the SAI's various legal missions. An evaluation involves the use of human and financial resources, so care should be taken to ensure that external evaluation requests accepted by the SAI do not compromise the mobilization of resources that would otherwise remain available for other optional or mandatory missions.

Finally, the conclusion of a feasibility study may be that it is not desirable to carry out the evaluation for the following reasons: effects that are too difficult to measure or of little importance, stakeholders that cannot be involved, other ongoing evaluation processes, insufficient qualitative or quantitative resources, etc. In this case, it may be desirable to reduce the initial ambitions and carry out a more traditional performance audit, which will measure the results of the policy in terms of the objectives defined and the resources allocated in order to monitor its effectiveness and efficiency.

Otherwise, the feasibility study should propose a statement of work, which will define the framework for the evaluation approach and validate the different elements (policy description, identification of stakeholders, evaluation questions, organization of the evaluation approach, selection of methodology and scientific instruments).

In particular, the statement of work should specify the evaluation questions and therefore define the questions that the evaluation will have to answer. The list of questions should focus on the expected results of the evaluation.

The questions formulated therefore have many reasons to be asked: to provide, through their answers, knowledge on the results and socio-economic impact of a policy, to inquire about the causal relationship between the measures in place and the measured results and to provide a judgment/estimate of the value of the measured effects and the defined objectives. The choice and formulation of questions should not be biased by the views of the different stakeholders, while at the same time trying to take them into account.

C.2 Evaluation Management

1) Organization

The SAI carrying out an evaluation, either by self-referral or at the request of an external party, should establish an organization comprising:

- A team of evaluators within the SAI. This evaluation team should include permanent members of the SAI or experts recruited by it. Since the technical capacity of the SAI cannot cover all areas, it is usually necessary to use experts in evaluation techniques and experts in the field covered by the evaluation to be carried out.
- A decision-making body. Under the full responsibility of the SAI, this body is responsible for:
 - validate the feasibility study and decide whether to proceed with the evaluation (in the case of a voluntary evaluation) or accept a request from a sponsor;
 - define the questions to be asked during the evaluation;
 - to schedule the evaluation and integrate it into the work of the SAI, allocating the necessary human and financial resources;
 - discuss and validate the provisional results and the final evaluation report;
 - to transmit, if applicable, the report to the sponsors;
 - to disseminate and promote the results.
- A monitoring body composed of members of the decision-making body and stakeholder representatives may also be established to bring together stakeholder representatives and allow them formally to monitor the work of the evaluation team throughout the process. This ensures that the work is monitored by stakeholders from the beginning of the process, collects the data requested of them, takes their views into account (questionnaires, interviews) and allows for dialogue on the interim and final reports. Such supervisory bodies are systematically set up in some SAIs, such as in France, while other SAIs involve stakeholders without going through them (Switzerland, Finland). In all cases, this supervisory body must remain consultative and must under no circumstances take decisions on the methodology or conclusions of the evaluation.

2) Instruments and methodology

To carry out their various tasks, SAIs use professional tools and methods that ensure the objectivity and reliability of the collection, processing and analysis of data and information. These factors are even more crucial in an evaluative approach.

An SAI involved in an evaluation must therefore have a wide range of instruments at its disposal to carry out an evaluation not only based on quantitative indicators and qualitative studies, as is the case for most performance audits, but also on comparative cohort studies, randomized controlled trials, statistical or social data, surveys and interviews.

An evaluative approach involves in particular:

(a) a review of the global scientific literature

(b) initialing or, failing that, international comparison, created by travel abroad or by interviews via the diplomatic network, or by subcontractors (experts, academics, etc.);

(c) quantitative or qualitative indicators on the effects of public policy assessed by the SAI. All current quantitative evaluation methods can be used and applied to existing databases: - statistics and probability, econometric analysis,

(d) comparative cohort studies to assess the socio-economic impact of a policy, with two different methods available. The objective of these methods is to create, either explicitly (control group) or implicitly (*ceteris paribus* analysis, estimation of differences in differences), a fictitious situation in which public policy does not exist, a counterfactual situation with which the current situation is compared. Through this comparison, it is possible, to a certain extent, to assess the outcome of the policy analyzed.

- the first is based on the comparison of a group that benefited from the measure with a control group (it being understood that these two groups are as identical as possible, that they are theoretically identical, with the difference that one benefited from the measure and the other not) (for example, two groups of sociologically and intellectually comparable students of which only one would have benefited from a mechanism to support academic performance). This is the method generally used.
- The second is based on a comparison between two groups before and after policy implementation and consists of estimating "gaps in differences". Here again, one group is impacted socio-economically by the measure and not the other. In fact, comparing the situations of beneficiaries and non-recipients as mentioned above may produce results influenced by the existence of selection effects or time trends that affect both groups. Seen from another angle, the comparison of their situation with what it was before the implementation of the reform (before/after estimator) is not more satisfactory, as many factors could explain the changes noted in the beneficiary's situation without having anything to do with the reform. The objective of estimating the difference in differences is to combine these two approaches. In the simplest cases, two groups are observed before and after the implementation of the measure to be evaluated. This is applied to only one group. The estimator then compares the changes in the situation of the beneficiary group and the control group before and after the implementation of the measure over the same period. The term "difference in differences" embodies the very essence of the method, with the evaluation focusing on the differences between the two groups and the respective changes (temporal differences) in their situation during the period:
 - a difference must eliminate systematic differences between the two groups;
 - the other difference must eliminate the temporal evolution, which is assumed to be identical for both groups in the absence of a measurement.

The estimator identifies the unique causal effect of the measure to be assessed, assuming that the situation of the two groups could have experienced the same changes in the absence of a measure.

(e) to carry out surveys to find out the opinions of the beneficiaries of the policy, as well as the opinions of the actors who have implemented it. These surveys provide information on the effects felt rather than on the objective effects, but they remain useful for evaluation, as the effects felt influence the behaviour of policy stakeholders. These opinions can be collected:

- by quantitative surveys addressed to a random and representative sample of the population (and not a sample "chosen" by the SAI or by the actors of the policy evaluated). A non-representative sample biases the information collected.
- by qualitative surveys: one-to-one or group interviews (of homogeneous or mixed groups, or first homogeneous and then mixed), forums, etc. This approach complements the quantitative measurement of opinions with random and representative samples: the questions asked and answers received are thus richer and help participants to reflect and formulate their opinions. Mixed group and forum dialogues are also rich in information.

(f) in addition, the SAI can also contribute to the evaluative approach with all the information it has already acquired about the policy being evaluated during its other audit and control engagements in the area covered, in particular in the context of performance audits of public bodies and programs affected by the policy being evaluated. This enrichment, which comes from the SAI's sometimes "intimate" knowledge of the field and the organizations concerned, constitutes an essential added value of the SAI.

Finally, this whole process of information and data collection only exists as part of the evaluation process. It is a particular feature of this approach in relation to performance and regularity audits.

Another characteristic is to measure the direct effects of a policy on beneficiaries as well as the indirect effects of the policy on society as a whole, or at least on a part of society that exceeds beneficiaries. The analysis of the socio-economic impacts of the policy goes beyond the analysis of its immediate outcomes or results. It is part of a global perspective and must identify temporal trends in the policy being evaluated. This diachronic requirement is one of the characteristics of the evaluation, which also makes it possible to put into perspective the value judgment that will be made. It reinforces the fairness of the assessment judgment and the confidence of stakeholders in the assessment process, highlighting the difference between assessment and conventional or control audit.

3) Identification of resources

The scientific and transdisciplinary nature of evaluation requires the use of experts from different fields: economists, statisticians, sociologists, experts in the main field of evaluation, etc.

With regard to methodology, the SAI may establish a permanent team of experts dedicated to its evaluation programs. In addition to the financial and legal experts usually present in the SAI, this team may include statisticians, economists with or without expertise in the public sector, or even sociologists.

The role of experts in these areas is to participate in all evaluations carried out by the SAI, to provide technical advice and to carry out some necessary technical work.

But due to a lack of resources, SAIs may also have to recruit one or more recognised external experts, either in the methodological field or in the field of the policy being evaluated, most often on an ad hoc and part-time basis. These needs must be identified in the feasibility study.

The procedures for selecting experts should meet the same requirements for professionalism, rigour and independence as those applied to SAI auditors. When this is made possible by the SAI's Statutes, it would also be beneficial to give experts the same investigative powers as SAI auditors, and to ask them to respect the same professional obligations (in particular in terms of confidentiality).

These experts must be voluntary, however, it is recommended that they be remunerated to ensure their independence, involvement and the quality of their interventions.

C.3 Finalization of the results

1) Review of results

The final evaluation report may be preceded by interim reports.

Interim reports can be used to interact with stakeholders through the evaluation oversight body.

The draft final report must be prepared by the evaluation team. It contains all the elements listed in the feasibility study as well as the conclusions of the evaluation itself:

- Summary of the objectives of the evaluation
- The initial questions
- Identification of stakeholders
- Scientific methods and instruments
- Sources of information collected
- The general context in which the policy evaluated was implemented, with its historical development and, if possible, relevant comparisons abroad
- Presentation of data from other SAI work in the area in question (including performance audits), in particular on the organization put in place to implement the policy and its actual costs
- The presentation of the data collected by the answers to the evaluation questions (results, measurement of effects/socio-economic impacts)
- The analysis and interpretation of these data, carefully separated from the presentation of the data, to avoid confusing findings and interpretations; this section contains the value judgment with respect to these results and the measured effects, and examines the causal relationships between the actions taken and the observed effects.
- Stakeholder opinions collected during the work
- Lessons and conclusions drawn from this work in recommending corrective measures to be provided; this section may contain an assessment of the overall relevance of the policy and actions evaluated.

The draft report is then submitted to the SAI's decision-making body, appointed for this purpose at the beginning of the process according to the specific procedures of each SAI. The decision-making body shall validate or amend the draft report

In this phase, it is the content of the report that is specific to the evaluation; not its form.

2) Contradictory phase

Once the results have been reviewed and the analyses adopted on a provisional basis, it is essential that the evaluation report adopted by the decision-making authority be discussed with the stakeholders of the policy being evaluated during an adversarial phase.

The circulation of these draft reports ensures that the SAI's provisional analyses and conclusions as well as the transcripts of stakeholders' opinions are accurate and known to all.

Stakeholder reactions can be collected in writing or orally (at hearings).

To allow for more feedback from stakeholders, the conventional audit procedure can be used to complement this contradictory and peer review process.

However, the systematic integration of all stakeholders (including representatives of the beneficiaries of the policy being evaluated) is one of the characteristics of the evaluation approach.

3) Finalization of the results

Once stakeholders have expressed their views, orally or in writing, the deciding body of the evaluation shall amend and validate the final evaluation report, in accordance with its procedures and under its sole responsibility;

Even if the evaluation was carried out at the request of an external body, including a Parliament, the SAI must make a personal and independent judgment on the findings, analyses, conclusions and recommendations of the evaluation.

In this final validation phase of the report, particular attention should be paid to the expression of the SAI's recommendations on actions to be taken after the evaluation in the area of the policy being evaluated. Due to their importance, most of this information is addressed to Parliament or the government.

These recommendations can for example:

- acknowledge the overall relevance of the policy and actions implemented
- acknowledge the overall relevance of the policy, but the inadequacy or ineffectiveness of the actions undertaken, taking into account the direct or indirect effects identified
- question the overall relevance of the policy and the coherence of its objectives, and make **recommendations for redirecting, suspending or withdrawing the policy.**

IV. Dissemination and use of evaluation results

A. Dissemination of the evaluation report

The dissemination of findings is one of the fundamental principles of the evaluation. The SAI should take this into account and publish the evaluation report. The final report is addressed to the stakeholders of the public policy being evaluated, the authors of the request (in the case of an evaluation following an external request) and the public at large.

In the event that the evaluation has been carried out in response to an external request, the decision whether or not to publish the report may be taken in consultation with the requesting authority of the evaluation to the SAI. Nevertheless, the principles of independence of the SAI and publicity given to the evaluation imply that the final decision should be left to the SAI, even in the event of disagreement with the requesting authority (This point should have previously been addressed in the discussions between the SAIs and the referral authority at the time the request for the evaluation was made).

The SAI should do everything in its power to ensure that the recipients and readers of the report understand the specificities of the evaluation approach. An educational method would be, for example, to explain the choices made in terms of methodology.

In particular, the sensitive nature of the SAI's assessment of the overall relevance of a policy requires particular care in drafting the draft assessment, to ensure that it is well understood and to avoid biased or exaggerated interpretations, particularly by the media.

Finally, it would be desirable for the SAI to report publicly on its public policy evaluation activities, for example in its annual report, as soon as it has embarked on this path and carried out some such evaluations. The pedagogy developed around this mission will gradually lead to a better understanding of it and will also contribute to creating a climate of trust with public actors and other stakeholders.

B. Use the findings of the evaluation

A public policy evaluation is intended to be used in the public decision-making process and monitored, to an even greater extent than all other SAI work.

In any case, it remains essential that the SAI does not bear direct responsibility for any monitoring decisions taken, in order to be able to maintain its independence on the one hand and, on the other hand, not be accused of interfering with the decision-making powers of public authorities.

Although the process does not necessarily involve using the findings of the evaluation, the SAI may monitor the process as an external observer and thus confirm its results. To this end, it may organise interviews on the implementation (or non-implementation) of its evaluation within an appropriate period of time, for example within 2 to 4 years. This monitoring is systematically carried out by some SAIs, such as the Swiss Federal Audit Office.