

The Commission of Experts is moreover sceptical as to whether EUIPO, which is only responsible for the registration of EU trade marks and designs but not for patents, can fulfil the tasks assigned to it within a narrow time frame. Instead of developing new competences at the EUIPO in the years to come, the Commission of Experts believes that it would be better to use the existing competences at the European Patent Office.

Finally Harnessing the Potential of Data

The Commission of Experts has repeatedly emphasized the importance of data for the R&I system in its reports.⁹⁰ Data is the basis for gaining knowledge and making decisions. It feeds into the development of innovative products and services. Last but not least, it is a key element of the digital transformation. The range of applications for data is broad and extends from controlling digital applications in agriculture (cf. chapter B1) to the development and application of artificial intelligence models (cf. chapter B4) and the evaluation of policy measures to enable the assessment and, if necessary, adaptation of measures on this basis (cf. chapter A2).

The Federal Government has recognized the importance of data and has planned or initiated various measures to improve the provision of and access

to data. The Data Strategy it adopted in August 2023 provides an overview of these measures and the timescales for their planned implementation.⁹¹ Some of the measures have already been implemented. For example, it is very welcome that the Health Data Use Act has now been passed, paving the way for the utilization of health data to improve diagnosis and treatment for patients. More measures still need to be taken and implemented, such as the establishment of an agricultural data space and the passing of the Research Data Act.

Progress in Data Provision and Use Essential

Numerous proposals exist for improving the infrastructure and access to publicly funded data and for linking individual datasets.⁹² These involve, among other things, necessary adjustments to existing legislation and its inconsistent interpretation in the Länder (including the General Data Protection Regulation – GDPR, the Federal Statistics Act and the Tax Statistics Act).

The Commission of Experts reiterates that considerable progress in the provision and use of data is essential in order to achieve progress also in the areas of data application and to achieve the digital transformation. Artificial intelligence as a key enabling technology is just one example, albeit a most important one.

A2 Determining the Causal Effects of Interventions

Many of the evaluation studies on measures of research and innovation policy (R&I policy) carried out on behalf of the Federal Government do not allow any conclusions to be drawn as to whether the developments observed can actually be attributed to the policy measures analyzed. The main reason for this is that evaluation studies often do not fulfil the methodological requirements for a causal analysis, not least because the prerequisites

for the appropriate use of suitable methods are not always met. The lack of knowledge about the impact of measures impedes systematic and evidence-based policy learning. The potential for learning from the evaluation of and experience with implemented measures and adapting future measures where necessary so that their objectives are better achieved remains largely unexploited.

The Federal Government is therefore called upon to systematically and comprehensively integrate causal analyses of the effects of measures into the deployment of R&I policy measures and thus ensure the conditions for the proper and professional implementation of these analyses and their usability in policy learning. To this end, it is necessary to include causal analyses in the specifications when tendering for evaluation studies, to improve the availability of data for the evaluating organizations and to publish all commissioned evaluation studies.

Causal Analyses Required for Research & Innovation Policy Learning

Causal analyses of the effects of policy measures are necessary to evaluate the effectiveness and efficiency of these measures and thus enable policy learning. Policy measures are effective if they achieve the objectives pursued and efficient if there are no more cost-effective ways of achieving the objectives pursued. Policy learning implies that measures are continued unchanged, readjusted or, if necessary, abandoned based on improved knowledge.

The Commission of Experts considers the potential of policy learning in R&I policy to be high for the following reasons:

- R&I policy measures are aimed at a broad range of different R&I activities – from basic research to applied research and application. The various types of R&I activities differ in terms of their starting points and their relevance to policy support. Their results are associated with a varying degree of uncertainty that changes over time.
- The extent to which an R&I policy measure is suitable for supporting certain R&I activities depends heavily on the respective context. The effectiveness of a measure can vary greatly depending on the sector or region.
- R&I policy increasingly aims to steer R&I activities in completely new directions and thus contribute to mastering the grand societal challenges. In doing so, R&I policy often cannot simply fall back on established measures but must adapt them or break completely new ground.

With this in mind, causal analyses of the effects of measures can make a significant contribution to adapting and further developing R&I policy measures and thus improving their impact. Causal analyses are part of ex-post evaluations, which examine and assess measures retrospectively. In addition to such ex-post evaluations, process evaluations are often also required as a means of gaining insights into the feasibility of measures while they are still in place and, if necessary, making adjustments at an early stage. This is particularly true in the context of a mission-oriented and transformative R&I policy, where numerous measures are implemented in parallel or sequentially, the interaction of which can only be incompletely assessed ex ante.

Informative Value of Impact Evaluations Depends on Methodology

The aim of causal analyses of the effects of measures is to determine whether observed changes in predefined target variables are causal, i. e. can be attributed to the policy measure under investigation. The validity of such causal analyses and thus the potential for policy learning depend on the methodology used. This has developed fundamentally in recent decades. Experimental and quasi-experimental methods of causal analysis are now firmly established standards of empirical research in economics and the social sciences.⁹³

The Commission of Experts already pointed out the importance of meaningful evaluation studies in its 2010 Annual Report and since then has repeatedly emphasized the need to carry out causal analyses of the effects of measures.⁹⁴ In addition, in 2013 the Scientific Advisory Board of the then Federal Ministry for Economic Affairs (BMWi) stressed the importance of causal analyses for the evaluation of economic policy measures in one of its reports. The methodological requirements it defined are still relevant – also for R&I policy measures.⁹⁵

The prerequisites for carrying out meaningful causal analyses must already be considered in the design of policy measures and created when the measures are implemented. It is essential to compare the situation arising from the implemented measure with the situation that would have arisen without the measure, the so-called counterfactual state (cf. box A 2-1). This comparison is made using a suitable evaluation design. The main requirements are that

those covered by a measure, the treatment group, are compared with a suitable control group of those not covered (cf. box A 2-1) and that the effectiveness of the measure is tested using suitable experimental or quasi-experimental methods (cf. box A 2-1).

The best possible conditions for a suitable composition of treatment and control groups are provided by genuine randomization of participation in the measure as part of policy experiments (cf. box A 2-1). There are numerous international examples for policy experiments.⁹⁶ In the case of a support measure, for example, true randomization means that the participants are selected at random. If the number of cases is sufficiently large, randomization can ensure that the group of beneficiaries and the control group of non-beneficiaries do not differ statistically in all observable and non-observable characteristics with the exception of the characteristic ‘supported’ or ‘not supported’.⁹⁷ This integrates the counterfactual situation, which is so important for a causal analysis, into the policy measure from the outset.

Randomization can often be combined with a pre-selection according to content-related, funding-relevant criteria associated with the measure.⁹⁸ For example, basic eligibility can be checked first. Randomized selection only takes place within the group of those eligible for funding, and possibly only within a subgroup.⁹⁹

Randomized policy experiments are often not feasible, for example because the necessary number of cases cannot be achieved. If this is foreseeable, then quasi-experimental methods of causal analysis should be considered from the outset (cf. box A 2-1). Their applicability and informative value can be improved by a suitable design of the measure and by collecting the necessary data.

Methodological Potential not Exhausted in Evaluation Practice

In recent years, the Federal Government has commissioned many studies to evaluate R&I policy measures.¹⁰⁰ The Commission of Experts examined the question of how many of these studies claim to demonstrate causal effects of the measures analyzed and to what extent this claim is fulfilled.

To this end, 81 publicly accessible evaluation studies from the areas of responsibility of the Federal Ministry of Education and Research (BMBF) and the Federal Ministry for Economic Affairs and Climate Action (BMWK), published between 2009 and 2023, were analyzed using a standardized analysis scheme.¹⁰¹ The results of this study show that a reliable causal analysis was often not carried out in the past, even in those evaluation studies that make statements on the impact of measures. The results also indicate that the conditions for causal analyses in the evaluation of policy measures need to be improved, for example in terms of data availability.¹⁰²

In 59 of the 81 evaluation studies analyzed, the results were interpreted causally – regardless of the method used (cf. figure A 2-2). Treatment and control groups were compared in 24 of these 59 studies. Only one of these studies conducted randomized policy experiments. Quasi-experimental methods were used in six studies. In seven studies, only matching methods were used (cf. box A 2-1), which have limited suitability for causal analyses. In the remaining ten of the 24 studies, no methods were used that allow a causal analysis of the effects of the measures. This means that only seven of the 81 evaluation studies analyzed were able to draw reliable conclusions as to whether the evaluated measures were effective. In 45 of the 81 evaluation studies, however, developments were interpreted as causal effects of measures, although the methodology used does not allow this conclusion to be drawn.

A total of 40 of the 59 studies with a causal interpretation of results attest to a positive effect of the evaluated R&I policy measures and 19 of the studies attest to at least a partially positive effect. As a result, all the studies analyzed that made an impact statement also found a positive impact. No study concludes that effects of the measures analyzed cannot be shown, or indeed finds negative effects.¹⁰³ This pattern differs significantly from the results of the impact analyses found in the scientific literature. These often conclude that the effectiveness of the measures analyzed cannot be shown.¹⁰⁴

The Commission of Experts notes that most of the evaluation studies analyzed do not meet the requirements for a meaningful causal analysis. Randomized policy experiments have hardly ever been conducted in the past. Methods of quasi-ex-

Box A2-1 Explanation of Basic Terms of Causal Analysis

Causal analysis: Causal analysis refers to statistical methods for determining cause-and-effect relationships between different variables. In policy analysis, causal analysis is used to determine whether and to what extent a certain intervention, such as an R&I policy measure (the independent variable), has a direct effect on the desired result (the dependent variable).

Counterfactual state: The counterfactual state describes the hypothetical situation an entity (e.g. a person, a group, a company, a region) affected by an intervention (the so-called treatment) would be in if it had not been exposed to the intervention.

Randomization: Randomization means that the participants are randomly assigned to two groups:

1. **Treatment group:** This group is subject to the intervention the effect of which is being analyzed.
2. **Control group:** This group is either not subject to any intervention or is subject to a different intervention than the one analyzed. The results of the control group are used to approximate the counterfactual state of the treatment group in order to measure the effect of the intervention or to compare the effect of the different interventions.

Randomization is intended to ensure that there are no systematic differences between the groups before the intervention. This means that all known and unknown variables that could influence the result are evenly distributed across both groups.

Experimental methods: Randomized experiments, often referred to as randomized controlled trials (RCTs), are used in research to test the effectiveness of an intervention under controlled environmental conditions. When appropriately designed and conducted, randomized experiments provide the strongest evidence for causal effects of an intervention.

Quasi-experimental methods: In contrast to experimental methods, quasi-experimental methods are not based on genuine randomization. Instead, constellations are analyzed where participation in an intervention was 'quasi-random' and could hardly be influenced by the participants. For example, in the case of interventions that were limited to individual regions, the possibility of participation is quasi-randomized by belonging to the region. Therefore, comparable units from other regions can be used as a control group. Subject to certain conditions, which differ depending on the method, causal effects of interventions can be estimated in this way.¹⁰⁵

Matching procedures: Matching procedures¹⁰⁶ are statistical techniques whereby, for each unit of observation in the treatment group, one or more units are found in the control group that are similar regarding certain observed variables. Matching procedures can help to reduce distortions caused by non-random participation in measures. In the context of causal analyses, they are used in addition to quasi-experimental methods. Matching alone does not solve the problem that participation in measures can be influenced by non-observable variables.

perimental causal analysis that go beyond mere matching have also only been used in isolated cases although a slightly positive trend can be observed here.¹⁰⁷ This means that policymakers lack the basis to learn from the (non-)effectiveness of previous measures and to increase the effectiveness of future R&I policy measures.

Multiple Reasons for the Lack of Significance of Evaluation Studies

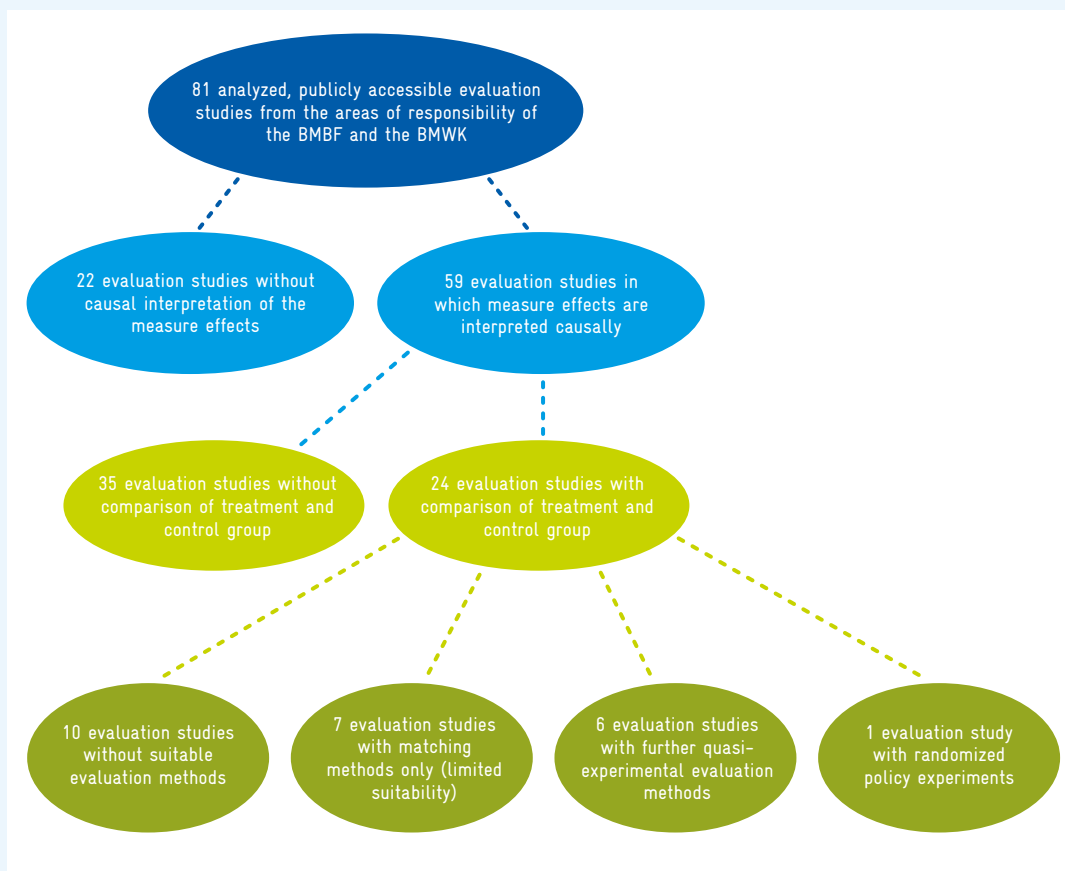
There are several reasons why the research design of an evaluation study may not or not fully meet the requirements for a meaningful impact measurement through causal analysis. Such reasons are cited in 32 of the 81 evaluation studies analyzed by the Commission of Experts. These reasons include:

- Insufficient data availability and quality make it difficult or impossible to identify suitable control groups and conduct meaningful causal analyses. Data that would in principle be available at the start of the measure are not always available for a causal analysis.
- Causal analyses can come up against methodological limitations. For example, a comparison of the treatment and control groups is not meaningful if the number of cases is too small. Moreover, it is not even possible to identify a control group if all potential recipients of a support measure are supported. It is also not possible to identify the effects of measures promptly if the effects of the evaluated policy measure can only be expected in the long term.
- The effectiveness of measures is evaluated based on defined targets. The prerequisite for this is that these targets can be measured or at least approximated. Unintended effects, both positive and negative, cannot always be measured or may be overlooked.
- Evaluating the efficiency of a measure is particularly demanding in terms of data, as this usually requires the effects of different measures to be measured and compared with each other. In many cases, however, at least a cost-benefit estimate can be made based on a causal analysis.

Fig. A2-2 Results on evaluation studies from the areas of responsibility of the BMBF and the BMWK



[Download Data](#)



Source: Own representation based on Büchele et al. (2024).
© EFI – Commission of Experts for Research and Innovation 2024.

Recommendations for Action

The Federal Government is pursuing ambitious objectives in the field of R&I policy and at the same time is subject to increasing cost-cutting constraints. A meaningful causal analysis of the effects of measures as part of the evaluation of policy measures creates the prerequisites for a learning R&I policy. The Commission of Experts therefore recommends improving the causal analysis of R&I policy measures as the basis for a learning R&I policy, building on the recommendations of previous annual reports.

Systematically Integrate Causal Analyses into R&I Policy Measures

- The conditions for a causal analysis of the effects of measures should already be considered when planning measures and the prerequisites for collecting the necessary data should be established. This means defining precise objectives and, if possible, specifying suitable indicators for examining the achievement of objectives.¹⁰⁸ Internal and external experts with proven expertise in the methodology of causal analyses should be involved already in this phase.
- When tendering for evaluation studies, the terms of reference should include a causal analysis and minimum requirements for the evaluation concept in this regard. A fundamental assessment of the efficiency of the measure, e.g. by means of a cost-benefit analysis, should also be sought. The timing of the evaluation should be chosen so that the relevant effects of the respective measure can be realistically observed.
- When commissioning evaluation studies, particular attention should be paid to the methodological expertise of the providers. Evaluations should not be carried out by organizations that are themselves responsible for the measure to be investigated. Suitable tendering procedures should be used to incentivize new qualified providers.¹⁰⁹

Optimally Align Causal Analyses with the Evaluation Context

- Randomized policy experiments offer ideal conditions for meaningful causal analyses. They should be utilized more often than before in R&I policy. Randomization is particularly suitable when an oversubscription of support measures is to be expected.
 - Regulatory sandboxes represent a special category of R&I policy measures that are intended to facilitate policy learning. When setting up regulatory sandboxes, the subsequent causal analysis should therefore also be considered, e.g. by ensuring that there are meaningful comparison units for the regulatory sandbox to allow its overall impact to be assessed.
 - When analyzing the interaction of different measures in the context of mission-oriented R&I policy, it can be useful to examine the achievement of the mission objectives through the interaction of the measures in aggregated form – possibly also in an international comparison.¹¹⁰
 - Not conducting a causal analysis of the measure effects may be justified in individual cases, not least because the conditions for this cannot be created in every case. However, this must be justified by the particularities of the respective measure context.
 - Where the requirements for robust causal analyses are not met, the use of simpler methods such as matching procedures can provide indications of the effectiveness of measures. However, in such cases, no statements should be made about the effectiveness of measures that are not covered by the methods used.
- ### Expand Expertise and Create Transparency
- The primary prerequisite for policy learning by means of causal analyses of the effects of measures is that they are conducted in an open-ended manner and that both positive and negative results are equally valued as advances in knowledge.¹¹¹

- Irrespective of the results, all analyses of policy effects commissioned by the ministries should be published. The data and programming codes used should also be disclosed or made available to interested researchers on request. This will enable the results to be replicated, possibly using improved methods of analysis developed at a later date.
- The institutional embedding of evaluation practice in the ministries should be further intensified and particular attention should be paid to the further training of staff in the units formulating policy measures.¹¹²

Increase Availability and Quality of Data

- The units or project executing agencies responsible for implementing a measure should systematically provide the evaluation teams with all documents associated with a measure. This includes, for example, draft proposals, information on all applicants and not just those who ultimately receive funding, as well as the preparatory documents and minutes of the selection and funding committees.
- The data laboratories in the ministries should be further developed into research data centres that collect and process data from all R&I policy funding measures and make it available for research.¹¹³ Uniform definitions and, where possible, indicators should be developed and used across ministries in order to ensure the comparability of data across different measures.
- Data that is already collected for administrative purposes is often of great importance for evaluation studies. Access to this data for the purpose of evaluating policy measures should be subject to a light-handed regulation and facilitated comprehensively, i. e. also at lower levels of data aggregation.¹¹⁴