# How to design and evaluate a Regulatory Experiment?

A Guide for Public Administrations

## Authors and contact details:

Prof. Dr. Dierk Bauknecht (**d.bauknecht@oeko.de**),
Dirk Arne Heyen, Dr. Peter Gailhofer
**Oeko-Institut e.V.**
Freiburg, Darmstadt, Berlin

Prof. Dr. Kilian Bizer (**bizer@wiwi.uni-goettingen.de**),
Dr. Daniel Feser
**Chair of Economic Policy and SME Research**
Faculty of Business and Economics
**Georg-August-University Goettingen**

Prof. Dr. Martin Führ (**martin.fuehr@h-da.de**),
Simon Winkler-Portmann
**Society for Institutional Analysis (sofia)**
at the **University of Applied Sciences Darmstadt**.

Prof. Dr. Kilian Bizer (**bizer@wiwi.uni-goettingen.de**),
Thore Sören Bischoff, M.A., Dr. Till Proeger
**Institute for Small Business Economics** at the
**Georg-August-University Goettingen** (ifh Göttingen)

SPONSORED BY THE

**Federal Ministry of Education and Research**

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

i f h
Volkswirtschaftliches Institut für Mittelstand und Handwerk an der Universität Göttingen

s o f i a
Sonderforschungsgruppe
Institutionenanalyse

**Öko-Institut e.V.**
Institut für angewandte Ökologie
Institute for Applied Ecology

h_da
HOCHSCHULE DARMSTADT
UNIVERSITY OF APPLIED SCIENCES
fbgw
FACULTY OF SOCIAL SCIENCES

# Table of contents

# Executive Summary

Experimenting with regulations? What may sound strange at first can be a fruitful way to support learning processes on technical, social or regulatory innovations and better assess their impacts ex-ante.

## What are regulatory experiments?

Broadly speaking, **regulatory experiments can be defined as a means to deliberately deviate from the current regulatory framework to try out new or different rules in a real-world setting.** General characteristics are the key role of public regulations, the involvement of government actors (whether local or national), and the generation of learning processes.

When speaking of "regulation" and "regulatory experiments", we mean the **whole range of public policy instruments, procedures and organizational structures**.

We differentiate between two key types of regulatory experiments (see text box, and Chapter 2 for details). While in type 1 – "Regulatory Sandboxes" – regulation is mainly the framework of socio-technical experiments, in type 2 – "Regulatory Innovation Trials" – regulation itself is the main object of experimentation and learning.

→ **"Regulatory Sandboxes" aim to allow testing technical, social or organizational innovations by (initially) temporary exemptions from existing legal rules** (e.g. through experimentation clauses). Examples range from exemption clauses for testing autonomous driving and delivery to regulatory sandboxes for testing innovations in the energy and fintech sectors.

→ **"Regulatory Innovation Trials" (RITs) aim to test new regulatory options and learn about their impact before introducing them on a permanent basis and eventually nationwide or elsewhere.** Examples range from basic income experiments with a limited number of participants to testing new traffic rules in a geographically limited area.

The distinction between the two types is ideal-typical: Like RITs, Regulatory Sandboxes can also entail experimental adaptations of subordinate regulations and represent a new legal situation. In turn, RITs also concern technical artifacts, infrastructures or social practices as part of the real-world setting in which they take place. However, the distinction is still useful since the focus and learning goals of the two types of regulatory experiments differ as described in the text box (see Chapter 2 for details).

## What are such experiments good for?

As the Council of the European Union (EU) concluded in 2020, regulatory experiments "provide the opportunity for advancing regulation through proactive regulatory learning, enabling regulators to gain better regulatory knowledge and to find the best means to regulate innovations based on real-world evidence, especially at a very early stage, which can be particularly important in the face of high uncertainty and disruptive challenges, as well as when preparing new policies" (Council document 13026/20).

What the Council of the EU concluded mainly regarding Regulatory Sandboxes is even more true for RITs. Being standard in the case of technical innovations in R&D processes, testing innovative regulatory options in a small but real-world setting before being "rolled out" can **help to better design effective and efficient regulation from the outset**. Addressees of regulation but also society as a whole can benefit from a less costly but still goal-effective policy design.

Regulatory experiments seem particularly useful for **addressing the major sustainability challenges** that our world currently faces. Any kind of governance dealing with these challenges and potential solutions must cope with profound problem interdependencies and uncertainties about the effects of interventions. This calls for a "culture of analysis" in political decision-making, thoroughly scrutinizing the benefits and burdens related to a regulation. The knowledge gained in experiments is supposed to complement other methods for ex-ante policy-impact assessments such as modeling.

At the same time, regulatory experiments should not be used without a serious intention to learn, i.e. neither to justify policy choices already taken nor to delay policy decisions for tactical reasons.

**Aim & background of these guidelines**

The following guidelines aim to **support officials in public administration** (whether in the EU, national ministries, agencies, or local administration) **in preparing, implementing and/or evaluating real-world regulatory experiments**. The overall aim is to foster regulatory learning and the development of a better regulation around societal efforts towards the UN Sustainable Development Goals (SDGs).

Public officials who are both experienced and inexperienced in regulatory experiments should benefit from the recommendations and best practices presented. The guidelines are written in a generic manner, primarily addressing the main actor in charge who drives and oversees the process – although different actors, different administrative units or organizations might be responsible for different tasks and steps.

The regulatory experiment may be located in different contexts and different phases of the development of new regulation. For example, the experiment may be carried out at an early stage to test new regulatory options, or it may be part of a more formal impact assessment, when certain options have already been pre-selected in the policy process. Experiments may be stand-alone or part of a larger program. These guidelines are written to be useful for these different kinds of experiments.

The guidelines are based on insights from a research project on regulatory experiments that included an extensive literature review as well as a detailed analysis of 27 cases of regulatory experiments (see Annex) from all over the world related to Sustainable Development (see text box on page 2 for further information on the project).

## Key messages & recommendations

➥ Regulatory experiments <u>can help policymakers to design more effective and efficient regulation</u> by gaining field knowledge of the benefits, costs and side effects. Such real-world experiments complement other methods for ex-ante policy-impact assessments.

➥ While <u>"Regulatory Sandboxes" primarily allow testing socio-technical innovations</u> by granting exemptions from certain legal requirements (see Chapter 2.1), <u>RITs allow testing regulatory innovations</u> before being rolled out (see Chapter 2.2).

➥ Given that experiments are about learning, a "culture of analysis" with an <u>open-minded attitude, a proper setup of the experiment, continuous monitoring and stringent evaluation</u> are key to maximize knowledge gains from the experiment (see Chapters 3.1 and 4).

➥ <u>Broad stakeholder involvement</u> – including all actors potentially affected by the socio-technical or regulatory innovation – is another key factor for success (see Chapters 3.2 and 4.2.3).

➥ Regulatory experiments entail several <u>legal issues</u> that need attention, from the overall legal framework conditions to the legal base of the specific experiment (see Chapter 3.3).

➥ While there is no standard procedure for regulatory experiments, the <u>ideal-type phase approach in</u> these guidelines (see Chapter 4) can help to properly design, implement and evaluate such experiments.

# 1  Regulatory experiments: enhancing field knowledge for better regulation towards Sustainable Development

**Designing effective and efficient public policies and regulations from scratch or adapting them to new developments has always been a challenge.** Unintended side effects may emerge, the willingness of actors to contribute to regulatory goals may be low, and the complexity of envisaged interaction might have been underestimated. Today's intertwined world with many multi-level and multi-actor interdependencies, major societal challenges – as captured notably by the United Nation´s Sustainable Development Goals (SDGs) – and trade-offs between goals make policy design a highly demanding challenge.

Typical ex-ante policy impact assessments usually rely on models with many assumptions under uncertainty, and past experiences. However, the real effects of a changed regulatory framework often differ because – inter alia – reactions of individual or corporate actors to the new regulation do not follow the (often linear) expectations. Technical, social and organizational innovations that might serve as solutions to sustainability problems face regulatory challenges like legal barriers and a wide range of uncertainties. This might be the case simply because the legal status quo could not anticipate current developments. Digital technologies have increased the gap between business innovations and regulatory timeframes.

**As part of responsive governance that aims to adequately cope with new developments, regulatory experiments can be an effective approach to deal with the aforementioned challenges.** As the Council of the EU concluded in 2020, "flexibility and experimentation can be important elements for an agile, innovation-friendly, future-proof, evidence-based and resilient regulatory framework" (Council document 13026/20, p. 3). For public administration officials responsible for policy design, in particular regulatory experiments "provide the opportunity for advancing regulation through proactive regulatory learning, enabling regulators to gain better regulatory knowledge and to find the best means to regulate innovations based on real-world evidence, especially at a very early stage, which can be particularly important in the face of high uncertainty and disruptive challenges, as well as when preparing new policies" (ibid., p. 5).

In the following, we clarify the term "regulatory experiment" and differentiate between two ideal types, presenting use case and real-world examples for each of them (Chapter 2). Afterwards, key issues around experiment design and implementation challenges are discussed (Chapter 3). Although not intended as a deterministic step-by-step playbook, the main chapter (4) presents concrete recommendations along different phases of an experiment. The recommendations are again illustrated with examples from real cases.
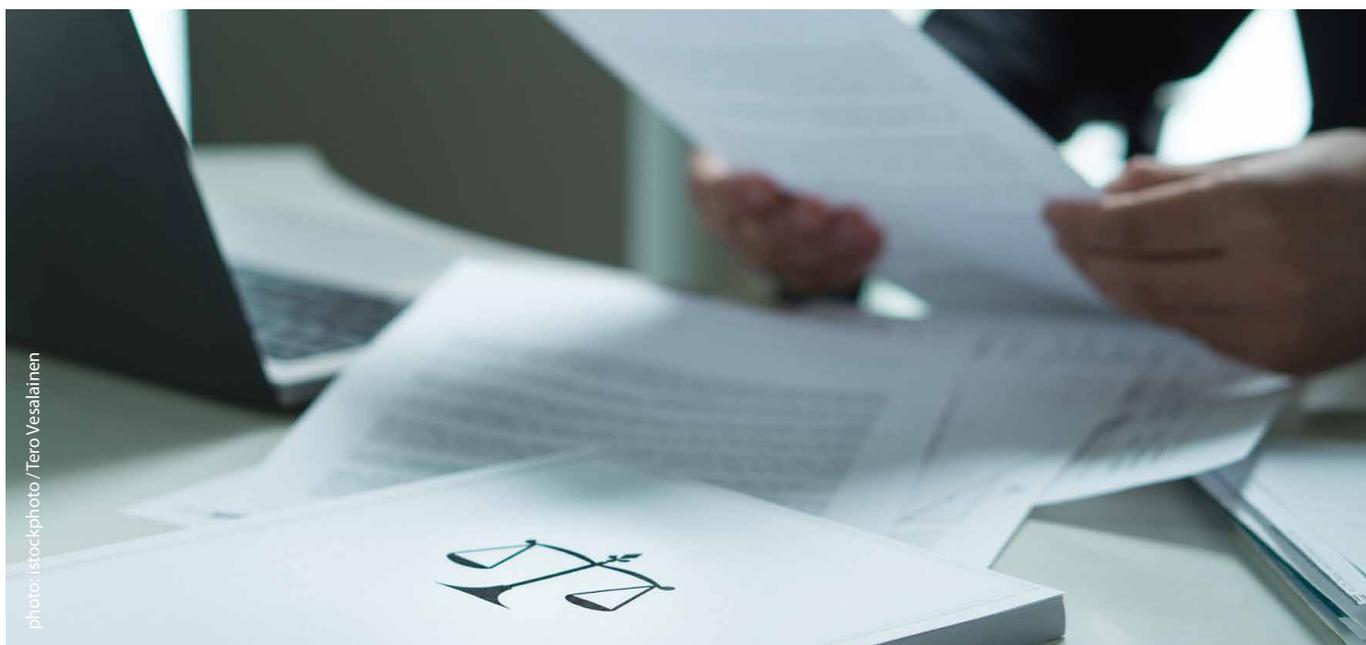


photo: istockphoto / Tero Vesalainen

# 2 Types of regulatory experiments and their respective benefits

In general, **regulatory experiments can be broadly defined as a means to deliberately deviate from the current regulatory framework to try out new or different rules in a real-world setting.** The main characteristics of such experiments are:

- the key role of public policies and regulations;
- the involvement of government actors (whether local, national or supranational); and
- the generation of learning processes.

When speaking of "regulation" and "regulatory experiments", in the following we mean not only prescriptive law in terms of "command and control" approaches or the narrowly-defined regulation of networks and monopolies (e.g. energy grids), but rather the **whole range of institutional arrangements of public policy instruments, procedures and organizational structures**.

We differentiate between two key ideal types of regulatory experiments. In the case of "Regulatory Sandboxes", regulation is mainly the framework of socio-technical experiments that may need exemptions from legal rules. In the case of "Regulatory Innovation Trials", regulation itself is the main object of experimentation and learning. The two forms are further elaborated in the following.

## 2.1 Regulatory Sandboxes: exemptions from existing legal rules

**Regulatory sandboxes aim to allow testing specific technical, social or organizational innovations for which the current legal framework poses significant challenges.** This if often the case with radically new technologies and business models, like in the context of digitization and automation.

The sandboxes are **characterized by (initially) temporary exemptions from the existing legal framework**. The exemptions or adaptations remove direct legal barriers (part of the innovation is not allowed) or economic barriers, i.e. the experiment is not economically viable under the current regulatory framework.

The Council of the EU defines regulatory sandboxes "as concrete frameworks which, by providing a structured context for experimentation, enable where appropriate in a real-world environment the testing of innovative technologies, products, services or approaches […] for a limited time and in a limited part of a sector or area under regulatory supervision ensuring that appropriate safeguards are in place" (Council document 13026/20, p. 4).

Specifically, regulatory sandboxes can be based on:

- exemptions from prohibitions;
- exemptions from or adaptations to prescriptive rules such as specific approval or documentation requirements, technical standards, or traffic law rules;
- adaptations to public tax or fee provisions;
- compensation of costs that would occur under the current regulatory framework.

Such exemptions or adaptations are usually based on public ordinances, which themselves are based on either a law specifically dealing with regulatory sandboxes, or an "experimentation clause"/"flexibility clause" in – for example – mobility law (see text box). Such clauses explicitly authorize a government to deviate from the existing law by a predefined degree.

> **Example of a general experimentation clause:**
> - Based on a general experimentation clause in the German Public Transport Act (Personenbeförderungsgesetz, PBefG), exemptions have been approved, for example, for testing autonomous driving and delivery as well as new forms and business models of car-/ride-sharing. The clause states: "For the purpose of practically testing new types or means of transport, the authorizing authority may, upon application in individual cases, approve deviations from provisions of this Act or from provisions issued on the basis of this Act for a period not exceeding four years, provided that public transport interests are not opposed thereto" (§ 2 Abs. 7 PBefG, own translation).

The regulatory sandbox concept originates from financial sector regulation in the UK, where it has also been adopted in other domains like the energy sector. Today, there are regulatory sandboxes in many sectors and countries, e.g. the financial sector in Denmark and Singapore, the energy sector in Germany and Norway (see text box), and the mobility sector – especially around autonomous driving and delivery

– in several European countries but also California, for example.

The regulatory exemption or adaptation of norms is usually not in the focus of an evaluation of the experiment, although one could assess its cost-effectiveness, for example. A proper evaluation of the experiment may lead to insights into whether and which longer-term policy changes would be appropriate. However, in order to determine optimal regulation around an innovation, one should switch to the second type of regulatory experiments (yet not limited to regulation around socio-technical innovations), as presented in the next section.

## 2.2 Regulatory Innovation Trials (RITs): experimenting with new regulatory options

In contrast to Regulatory Sandboxes focusing on socio-technical innovations (see previous section), regulation itself is the core of RITs. They are about **testing new or substantially-modified policy instruments ("regulatory innovations") in a real-world setting** to find the optimal regulation to achieve certain political goals.

Being standard in the case of technical innovations, testing regulatory innovations before being "rolled out" can help to design better regulation from the outset. When experimenting with new regulations, one might discover that they do not work as intended, e.g. an emissions trading regime might not sufficiently incentivize companies to reach emission reduction goals, or implementation costs may be higher than excepted. This represents valuable "regulatory learning" and can help to avoid potential costs and downsides of introducing inefficient or otherwise problematic regulations nationwide.

The **evaluation of regulatory options is at the core of these projects**, regarding such criteria as effectiveness, efficiency, justice implications, acceptance and unintended side effects. The knowledge gained from real-world experiments can complement other ex-ante policy impact assessment methods like modeling. RITs should not be used without a serious intention to learn, i.e. neither to justify policy choices already taken nor to delay policy decisions for tactical reasons.

To compare different policy options, public authorities may even set up several parallel RIT in different regions or city quarters, or sequential RITs in the same community.

Examples of RITs range from local to national governance levels.

To obtain substantial evaluation results from experiments, (natural) scientists often use **Randomized Control Trials**. Such experiments randomly allocate individuals to two or more groups, where some receive the intervention studied (treatment group) and some do not (control group). Researchers then compare the two groups with respect to a measured response of interest. By relying on randomization and only varying one factor at a time, using a Randomized Control Trail design significantly increases the chances of producing valid cause-effect relationships.

Randomized control-like experiments can also be used to study targeted policy interventions, as the Finnish basic income case (s. Textbox above) shows. The group receiving the basic income ("treatment") was similar to the rest of unemployed people in Finland ("control group") in all relevant background characteristics. However, controlling potentially intervening variables is a challenging task in real-world settings and only possible to a certain degree. In the Finnish case, the evaluation was further complicated by a general policy reform of entitlement criteria for unemployment benefits in 2018.

Nevertheless, learning can also occur without explicitly randomizing subjects into treatment and control groups, e.g. learning about practical implementation issues, stakeholder reactions, and politics. In these cases, the interaction between different actors is important to facilitate learning. What all forms of experimentation have in common is that monitoring and evaluation is important to maximize knowledge gains from the experiment.

photo: istockphoto / skyNext
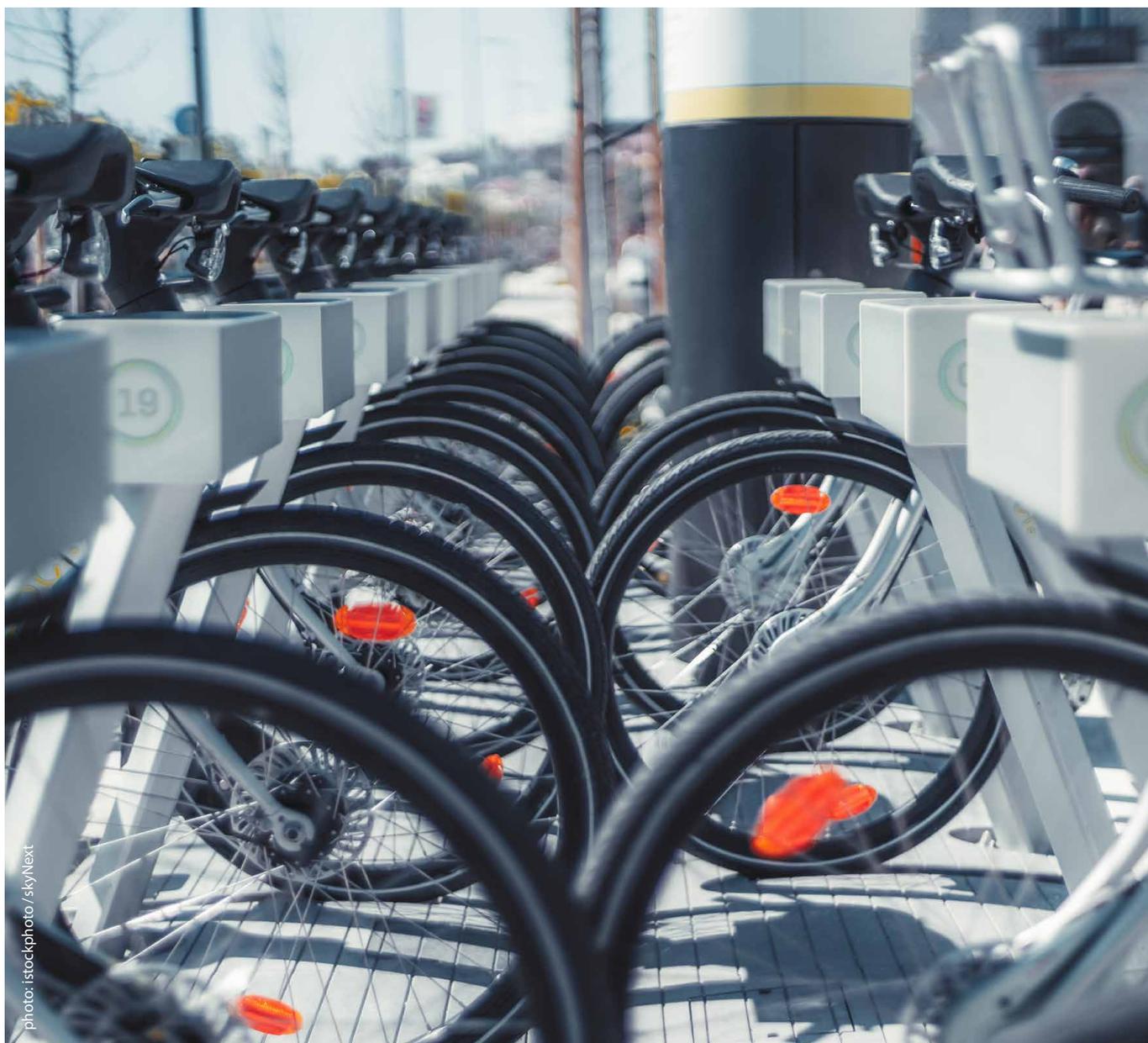
# 3   Key issues to consider for regulatory experiments

Here, we present key issues to consider when designing and implementing regulatory experiments.

## 3.1 Learning & evaluation throughout the experiment

**Experiments are about learning:** It is important to conduct experiments with this objective in mind. The experimental design should enable as much learning as possible. One can learn about the effects of the instrument on various stakeholders, as well as learning about processes, practical implementation issues, stakeholder reactions, and politics. A key question – even before the experiment is set up – is whether an experiment is the most appropriate tool for learning in the specific case (see Chapter 4.1).

> **Type of regulatory experiment**
> Learning is particularly relevant in the case of RITs, which are explicitly set up to learn about new regulatory options. Nonetheless, it should not be neglected in the case of regulatory sandboxes either, even though these experiments focus on non-regulatory innovations. Even in these cases, it is beneficial to draw up the regulatory exemptions and the design of the experiment in a way that allows for as much regulatory learning as possible. Regulatory sandboxes with regulatory exemptions that have no chance of being implemented in the future are therefore not recommended.

**Learning throughout the experiment:** Learning is not something that starts once the experiment is finished and evaluated. While an explicitly planned evaluation after the experiment is highly important, learning should start well before that and play a key role throughout the experiment. Indeed, how successful learning can be already depends on the design of the experiment. It should be clear from the outset which questions the experiment is supposed to answer, and the design of the experiment needs to be geared towards answering these questions.

For example, it needs to be decided at the very beginning whether different policy options are to be tested in parallel, and if randomization and control groups are needed. How should results be dealt with that may be biased through the experimental setting and the limited timeframe of the experiment? How can long-term effects of the tested instrument beyond the duration of the experiment be anticipated, and how does the tested instrument affect investment decisions that remain relevant even when the experiment is finished? Moreover, how can unintended consequences of the policy instrument be dealt with within the experiment?

If these issues are not adequately reflected in the initial design of the experiment, they can hardly be fixed at a later stage and will thus weaken the validity of any evaluation of the experiment.

It should also be an explicit option that the experiment fails and that the tested instrument is not useful. This can be an important learning result. Moreover, this option can help to address concerns that the experiment prejudges future regulation.

Learning should also be a guiding principle in the course of the experiment. This also helps to deal with the uncertainty that is inherently part of experiments. Such continuous learning helps to better understand the tested policy instrument. It can also be important to fine-tune the experiment in the process and adapt it to new insights. Experiments are also about learning to learn.

When experiments are set up to test future regulatory options, it is highly recommended to **develop an understanding of what the socio-technical context could look like in the future,** which regulatory challenges are likely to arise, and which regulatory options may become relevant. This is particularly important in the context of sustainable development, where dynamic changes can be expected. Experiments need to be coordinated with expectations and emerging requirements, i.e. the future context in which the tested policy instrument needs to function.

> **Type of regulatory experiment**
> This also points to a major difference between exemptions from existing legal rules (Regulatory Sandbox) and testing new regulatory options (Regulatory Innovation Trial). Learning ideally requires testing new solutions, and not simply an exemption from existing solutions.

**Upscaling and transferring the results:** Importantly, learning does not only refer to the experiment as such. The experiment ultimately aims to improve the policy instrument and apply it in a broader context. This means upscaling it from the restricted experiment to e.g. the national policy level, as well as transferring it to other contexts. The learning design of the experiment should be in line with this objec-

tive. Again, experiments should consider this from the very beginning.

In terms of upscaling, the experimental design ideally also includes a procedure concerning how to deal with the learning results of the experiment: Does the experiment simply finish, or is there a decision procedure for how to take the results further?

**Overall, when conducting experiments, one should invest in a robust methodological design and useful learning conditions, even in the face of higher resource requirements and potential resistance.**

## 3.2 Stakeholder involvement

The outcome of an experiment depends on who participates. While experimenting with new regulation primarily affects the regulator and the regulated, the group of actors who should be involved in an experiment typically needs to be broader.

> **Type of regulatory experiment**
>
> **This is the case especially in Regulatory Innovation Trials, where the effect of new regulation on various stakeholders is relevant. If the experiment is more about regulatory exemptions to test technical innovations, this may be less critical from a regulatory perspective. However, it is still important to include the actors who are affected by this innovation.**

**Appropriate participation needs to be organized throughout the experiment.** As with learning, participation is not something that happens automatically or "on the side", but rather it requires explicit attention, appropriate tools and instruments to foster exchange between participants. Experiments are about learning, but also about politics. Broad stakeholder involvement is relevant for both.

If only the regulator and regulated take part, other stakeholders will not learn about the new regulation, and it will be difficult to ascertain how it affects them. Broad stakeholder involvement is one way to deal with the unintended consequences of a new instrument. Moreover, a heterogeneous set of participants may be a prerequisite for drawing universally applicable conclusions about the effect of the regulatory options.

An active role of the regulator in the experiment cannot be taken for granted, as regulators are used to regulating, but not to experimenting. Nonetheless, regulators should consider experiments as a useful tool for themselves to learn more about regulatory options. Indeed, especially if the experiment also includes the question of how a new instrument can be implemented in practice, regulators should not only be observers, but rather they should become directly involved (learning by doing).

An involvement of regulators or law makers is relevant for another reason, namely involving those who decide about how to use the results and whether they are applied more broadly (upscaling) after the experiment can help to increase the impact of the experiment.

This leads to the **politics of experiments**. Experiments are not only about creating new knowledge, but also about developing support and an actor network for new regulatory options. Especially if new ideas are tested, experiments can help to build up public support. This requires broad participation and is also a prerequisite for learning. The involvement of key actors in relevant networks can support the recruitment of participants and distribution of information.

For the actors involved in designing, steering and evaluating the experiment an open-minded attitude, based on a "culture of analysis" is key. This needs to be defended against any attempts to "capture" the process due to political preferences; a dispute that might occur at any time.

## 3.3 Legal issues

From a legal perspective, regulatory experiments face **four general questions**:

- First, are exemptions from the given legal framework necessary to carry out the experiment?
- Second, does the experiment itself require the enactment of new legal regulations?
- Third, which authority has the legal competence to issue the relevant provisions to permit and/or perform the experiment?
- Fourth, are there any potentially colliding legal rules or principles – apart from the norms to be potentially substituted as a part of the experiment – that might prevent the experiment or lead to unwanted legal consequences?

Although the answers may diverge, these general questions are relevant for both types of regulatory experiments explicated in Chapter 2. The necessity of **exemptions from existing legal rules** is given – by definition – in regulatory sandboxes. In these cases, specific given rules or standards have been identified as obstacles for social or technological

innovations and shall be removed for specific cases during a limited time to permit experimentation. However, the question concerning the way in which the adaptation of the existing legal practice can be performed can be answered in different ways: in the regulation in question, experimentation clauses or other legal options that allow an experimental deviation from the given legal practice may already exist. Only if such leeway does not yet exist is an exemption from the relevant regulatory statutes necessary, with the consequence that the competent institutions have to adapt these norms or issue new regulations in accordance with the specified procedures.

In the case of RITs, this question is slightly more complex. On the one hand, such a regulatory experiment itself comprises the modification or enactment of norms, which may or may not be statutory norms. On the other hand, issuing such an experimental norm again can require an adaptation of superior law: just as in the case of regulatory sandboxes, superior law can impede experiments with innovative rules if it does not provide the relevant leeway for deviation from the given legal practice. An experiment with innovative rules in a field of law that already is densely regulated thus also requires adapting the respective laws by the competent legislator, e.g. by means of establishing experimentation clauses.

Even if no colliding laws exist, testing new or substantially-modified policy instruments may presume **enacting new legal statutes** to allow for the RIT. Rule of law principles can presuppose that the administration has a legal basis for its activities, in particular if these activities affect fundamental rights or hold relevance for other essential issues of the common good. RITs that aim at testing policy instruments that do not affect fundamental rights and other important policies therefore may be possible without a specific legal basis. For example, the existing leeway for legal experiments can be wider with respect to changing existing policies regarding subsidies or changes of public institutions and governance mechanisms. Thus, it can be possible to perform regulatory experiments without a specific legal authorization. However, in many cases, regulatory experiments will require a specific legal basis. In principle, this will be the case when new regulations are supposed to comprise a legal mechanism to ensure compliance. Such a necessity has been observed with respect to different legal systems: For example, there was no law at the national or regional level for a regional emission trading pilots in China, only an ordinance by the central government. This later proved to be a problem when regional governments had to design compliance systems for the experiment, as possible fines had no legal basis due to a regulatory gap in the ordinance.

According to the aforementioned third question, in cases in which the adaptation of the legal framework is necessary to carry out a regulatory experiment, it has to be clarified which legal entity has the **legal competence to issue the relevant provisions** to permit and/or perform the experiment. Competences rules differ from country to country.

### Examples of competences concerning experimental legislation in Germany

Experimental legislation at the federal level in Germany in accordance with the Fundamental Law (i.e. the German Constitution) is conceivable – for example – in the area of air transport (possibly for drones), (autonomous) railways, commercial law and road traffic (autonomous driving). On the federal level, the "Länder" may also enact experimental legislation for themselves, as well as experimentation clauses to deviate from the standards provided for in their law. Relevant legislation of the state level already exists. For example, there are experimental clauses under municipal law to promote possible reforms/innovations at the local level (Wittig et. al. 2020). Finally, experimental legislation can – under certain conditions – also be enacted by the government: in Germany, in accordance with Article 80 (1) of the Basic Law, the Federal Government, the Federal Ministers and the Governments of the States may issue statutory ordinances, if provided for in a federal law.

As a final general point, experimenters also must consider **rules and principles outside of the experimental scope**. For example, this applies to the fundamental principle of equality, which may become relevant when people are confronted with different legal or procedural regulations due to spatially limited experimental regulations. The principle of equality also is concerned when regulatory experiments only give individual companies the opportunity to test certain innovations. However, as has been found for the German context, legal objections regarding unequal treatment may be rebutted if experimental clauses are substantiated with foresight and designed in such a way that decisions at the administrative level are comprehensible in terms of both the substance and the procedure. (See the following publication: Holger Schmitz, Christian Alexander Mayer, Carl-Wendelin Neubert, Ines Reiling, Umsetzung der BMWi-Strategie „Reallabore als Testräume für Innovation und Regulierung": Erstellung einer Arbeitshilfe zur Formulierung von Experimentierklauseln (Los 1), BMWi., 2020)

## Large variety of legal issues in concrete cases

Besides such general issues, the formulation of an experimentation clause can lead to a variety of particular questions. For example, liability risks should be clarified in cases in which experiments involve testing real-life applications or technologies that could affect the legal interests of the involved enterprises or third parties. Regulatory experiments should thus systematically consider such legal issues. Recent research has therefore developed detailed guidelines to ensure that such clauses can be formulated in a permissible and practicable manner (cf. Schmitz et. al., 2020). Such questions also became relevant in the context of the experiments studied in our project: the Finnish basic income experiment had to address contradictions to the Finnish constitution, in addition to possible conflicts with other national legislation, e.g. social laws and European legislation. In the Dutch electricity sandbox, the European consumer protection framework and the Dutch tax act had to be taken into account. Similarly, the Italian electricity sector experiments had to consider the market anti-discrimination and consumer protection legislations when designing derogations.



photo: istockphoto / sommaiphoto

# 4   Design, implementation and evaluation of regulatory experiments

This chapter provides detailed step-by-step options for designing and implementing regulatory experiments. Experiment can be divided into four ideal-typical phases.

While an experiment is a fixed-term project with a clear start and end (as depicted above), ideally the experiment is part of a broader circular learning process where results are fed back into the policy process, which then lea ds to new experiments.

The following table shows the key tasks for each of these steps, which will be explained in more detail in the following sections.

| Clarification: Goal and type | Preparation | Implementation | Evaluation |
|---|---|---|---|
| ◗ Is an experiment needed at all? <br> ◗ Clarify: goal of the experiment <br> ◗ Appropriate type: Regulatory Innovation Trial or sandbox? <br> ◗ Is there uncertainty concerning the future system and its regulation? <br> ◗ Define clear regulatory question | ◗ Embed experiment in broader context <br> ◗ Obtain political support, sufficient budget and clarify legal issues <br> ◗ Organize stakeholder involvement <br> ◗ Determine the concrete experimental project <br> ◗ Prepare evaluation and learning <br> ◗ Decide on size and duration of the project | ◗ Monitor the implementation <br> ◗ Be ready to fine-tune the experimental design <br> ◗ If needed, even adapt instrument design during the implementation <br> ◗ Manage stakeholder involvement | ◗ Evaluate the experiment to learn about <br> ◗ the tested regulatory innovation <br> ◗ how it can be upscaled <br> ◗ the experimental design <br> ◗ Recommend regulatory option(s) <br> ◗ Adapt regulatory option based on evaluation results <br> ◗ Prepare the upscaling of results |

## 4.1  Phase 1: Clarification: Goal of the experiment and determination of type

| Clarification: Goal and type |
|---|
| ◗ Is an experiment needed at all? <br> ◗ Clarify: goal of the experiment? <br> ◗ Appropriate type: Regulatory Innovation Trial or sandbox? <br> ◗ Is there uncertainty concerning the future system and its regulation? <br> ◗ Define clear regulatory question |

Regulatory experiments can be a powerful learning tool that should be applied more often. However, this is not to say that they should be applied in any case and without considering the specific context. Therefore, before a regulatory experiment is set up (see next phase "preparation"), the following questions need to be addressed:

1) Is an experiment needed at all, and of added value, in this case to meet the goal?

   a. Regulatory sandbox: Is there really a regulatory problem that prevents testing the socio-technical innovation? Or can it be tested within the existing regulation? It can help to take a closer look at what is possible within existing regulation before setting up an experiment.

   > **Case example**
   > **The UK Electricity and Gas Sandbox showed that affected businesses tend to be more in need of advice on possibilities within the existing legal framework rather than applying for a regulatory sandbox, as they are mostly unaware of what they are already legally allowed to do. Similar situations occurred within the Dutch energy sandbox.**
   > **(see Annex for list of cases)**

   b. Regulatory Innovation Trial: Do the benefits expected from the experiment exceed the costs? Costs not only refer to monetary costs but can also mean that the introduction of a new regulation may be delayed through the experiment. Are there knowledge gaps that can be expected to be closed through learning in an experiment? Or would it be possible to directly opt for a full-fledged implementation?

2) What should be the goal of the experiment and what type of experiment would be appropriate to meet that goal?

   a. Regulatory sandbox: Regulatory exemption to test socio-technical innovations

   b. Regulatory Innovation Trial: Testing a regulatory innovation

Depending on the type, the design of the experiment – which is further explained in the next phases – differs.

3) Before setting up a regulatory experiment, it is important to scrutinize whether there is really a regulatory problem that needs to be addressed to enable sustainable solutions for the future. Is there regulatory uncertainty concerning the transformation to a future system that meets the Sustainable Development Goals? In this case, a regulatory experiment can be particularly useful. Alternatively, **is there simply a request for a regulatory exemption to enable certain business models that are not viable otherwise** and where it is not clear why this business model should be promoted from a societal perspective? In the latter case, an experiment should typically not be set up.

4) As a result of this phase, there should be a clear regulatory question.

   > **Type of regulatory experiment**
   > **The regulatory question will be less pronounced in the case of regulatory sandboxes. However, even here, experimenters should have an idea about how they can benefit from the experiment for developing future regulation.**

Please note that this phase can bring about useful learning insights, even if an experiment is not started as a result.

## 4.2 Phase 2: Preparation of the experiment

### Preparation

- ➟ Embed experiment in broader context
- ➟ Obtain political support, sufficient budget and clarify legal issues
- ➟ Organize stakeholder involvement
- ➟ Determine the concrete experimental project
- ➟ Prepare evaluation and learning
- ➟ Decide on size and duration of the project

A thorough preparation of the experiment is highly important to maximize the learning effects that can be achieved. On the one hand, this includes measures regarding the context of the experiment like embedding the experiment in a broader learning context and getting political support. On the other hand, the setup of the experiment as such needs to be prepared.

Overall, the preparation of Regulatory Innovation Trials is more demanding than regulatory sandboxes to achieve the desired regulatory learning. Nonetheless, given that regulatory sandboxes can also generate regulatory learning, careful preparation is important in this case as well.

Previous experiments show **various complications that experimenters can face** when planning regulatory experiments, e.g. due to political conflicts, diverging interests of stakeholders and legal uncertainties. In the case studies analyzed, experimenters have addressed these complications by using various preliminary measures, e.g. by analyzing prior experiments to identify best practices or consulting scientific and political experts as well as stakeholders and the public. Sound preparation of the experiment has proven useful to determine what the objective of a specific regulatory experiment should be, how it should be designed for that purpose, and to secure the framework conditions for the experiment.

### 4.2.1 Embed experiments in a broader learning context

Regulatory experiments should not be stand-alone instruments but rather part of a broader portfolio of measures to improve regulation and facilitate regulatory learning. Experiments need to be embedded in this broader learning context, which includes:

- ➟ An analysis of the regulatory problem as well as the knowledge that is available to address the problem. This

should not be limited to current problems, but needs to include potential future issues, for which a regulatory option can be tested now.
- ➟ Complementary instruments that are available to fill the knowledge gap; for example, modeling studies, laboratory experiments, role plays.

Understanding the potential contribution of a regulatory experiment requires analyzing the regulatory problem. This in turn requires an understanding of why current regulation does not achieve the desired outputs and which alternative regulatory options are available.

This can be based – for example – on a delta analysis, including the following questions:

- ➟ What is the regulatory problem?
- ➟ What do we want to achieve with regulation (target state)?
- ➟ What prevents the actors in the current state from making contributions towards the target state? Why do the actors in the status quo not behave in the way that is required from the perspective of the regulator?
- ➟ What are the regulatory options to overcome this delta?

Before setting up the experiment, it is useful to gain an overview of the knowledge that is already available to answer these questions. It can also be helpful to fill specific knowledge gaps before the experiments is set up.

This includes analyzing the situation in the jurisdiction at hand, as well as gaining a good understanding of the experiences elsewhere with the regulatory instruments that are to be tested, or even the experiences with similar experiments that have already been carried out. All of this should be done anyway before regulation is changed, but it can also help to identify the specific question and setup of the regulatory experiment.

This approach shows that a regulatory experiment needs to be embedded in a broader process of generating regulatory

knowledge. An experiment can complement such a process, rather than replace it.

### 4.2.2 Obtain political support, sufficient funding and clarify the legal basis

Regulatory experiments require three key resources that should be secured in time: political support, a sufficient budget and a clear legal basis.

In general, the preparation of regulatory experiments is facilitated if all political actors involved in the legislative process have understood that there is a need for regulatory learning and consider regulatory experiments as an instrument to facilitate this. Consultations with other experts, political actors and the public may help to pinpoint different positions and interests towards the foreseen regulatory experiment (see also Chapter 4.2.3: How to involve stakeholders?). Coordination between and within governmental (municipal) departments and with regulatory funding agencies can also be highly relevant.

In order to increase public legitimacy, not only the goal and reasons but also **the costs of experimentation should be transparent**. There should be sufficient budget to cover all necessary steps as described in these guidelines, including the evaluation and upscaling phase; otherwise, the experiment may end without satisfying results, which could even be counter-productive.

Finally, a sound legal basis is a prerequisite. Deviating from the legal framework in place or enacting new, innovative rules is not trivial and frequently requires putting legal provisions in place. General issues to be addressed at an early stage concern the initial analysis of which legal stipulations may be subject to the regulatory experiment – i.e. should be suspended or enacted – and which legal body is competent to carry out the adjustments of the legal framework. A preliminary examination of more specific potential legal hurdles is also recommended at an early stage. Such legal issues can be numerous and diverse. Addressing these issues early and with the help of specific expertise can help to avoid legal complications at a later stage.

### 4.2.3 How to involve stakeholders

How to involve stakeholders requires explicit attention in the preparation phase, which includes considering the following issues:

1. Who should participate?

Explicit mapping of relevant stakeholders and their interest can help to answer this question.

The case studies have shown that having a heterogeneous group of participants is beneficial (or even a prerequisite) for learning from regulatory experiments. Nonetheless, there is also a trade-off between "leaving the door open" and targeting specific actors if the composition of participants needs to be representative.

2. How to build a broad consensus between stakeholders that the experiment is useful.

3. How to consider the knowledge and interests of stakeholders, and how to transfer learning results to stakeholders.

4. How does the role of stakeholders change during the experiment? In which phase should they participate, and in which way?

5. Which stakeholders should be on board at a very early stage to better understand their perspective and – at least informally – involve them in the design of the experimental space?

6. How to incentivize and motivate stakeholders.

7. How to organize concrete stakeholder involvement and exchange between stakeholders. There should be explicit formats for stakeholder participation and interaction.

> ### Case example
> **For a shared space pilot project ("Begegnungszone") in a street in Berlin, a "steering group" comprising representatives of local administration, associations, local businesses and a residents' initiative was established based on the suggestion of the latter. Their task was to discuss and agree on concrete steps for public participation around the project. Concrete participation formats included online and paper questionnaires as well as townhall meetings and planning workshops, The steering group also agreed to precede the permanent reconstruction of the street with a "test phase", namely an initial time-limited experiment as part of the locally-limited experiment.**
>
> **(see Annex for list of cases)**

Whether participation is compulsory or voluntary is a key question for experimenters; for example, in the case of

How to design and evaluate a Regulatory Experiment?
A Guide for Public Administrations

**17**

electricity customers and the regulation of electricity tariffs (see also the case example on the basic income experiments in Finland and Ontario in Chapter 4.2.5). If it is voluntary, participation may be distorted and only innovative customers or those who stand to benefit from the regulation may take part. If participation is compulsory, this may allow for more representative results. Nonetheless, experiments may need to ensure that customers in the experiment are not disadvantaged, which again may lead to a distortion of results.

Especially **if participation is voluntary, experiments largely rely on the motivation of the target group** to participate. Consequently, the experiment needs be to transparent, hurdles for participation need to be low and incentives like reimbursement of costs can be important. If there are high prequalification requirements or cumbersome administrative procedures, this may deter potential participants who might have been relevant for the overall success of the experiment. Experimenters might guide actors through the application process. If the new regulation has a realistic chance of being adopted after the experiment, this can also incentivize participation.

> ## Case example
> The experiments with voluntary participation have shown several measures to incentivize stakeholders to participate. The energy sandbox in Singapore provides a perspective for the time after the experiment: exemptions may be transferred to the legal framework so that the tested innovations can be introduced on a larger scale. The experimental framework integrates a decision procedure and pre-defines transition conditions for the end of the experiment.
>
> **(see Annex for list of cases)**

As for concrete methods for stakeholder involvement, the rich experience from transdisciplinary research projects can be used.

While a broad stakeholder involvement is important, the coordination of the core team should not be neglected. In previous experiments, the internal organization of the team that conducts the experiment including the interdisciplinary research team, intra-administration organization, and the coordination and setting up of processes between actors from the administration and research team have been pivotal. In particular, clarity of decision-making processes and the stringent application of formalized a-priori-agreed rules has been helpful to establish trust between the various actors.

## 4.2.4    How to determine the concrete experimental project

There are different ways to select the concrete experimental project and the specific question to be addressed. The selection process can be top-down or bottom-up, or a combination of both.

> ## Type of regulatory experiment
> In regulatory sandboxes, the selection process is rather bottom-up: innovators face certain regulatory barriers for which they request a regulatory exemption. However, in this case it is recommended that there is also a top-down process to assess whether the requested exemption fits into the "bigger picture" of where regulation should be going.
>
> In RITs, experiments should always start from a regulatory problem – and thus a systemic question – rather than a problem for individual innovations or business cases. Here, the regulator is the one who wants to ascertain how certain regulatory options work to address societal challenges. This regulatory question still needs to be developed with stakeholders, so a bottom-up approach should not be neglected here either.

The following table presents several options for how to determine the concrete experimental project, especially for RITs. **(Table 3)**

**Table 3:**   Options for determining the concrete experimental project, especially regulatory innovation trials

|  | How is the regulatory experiment determined? | Comments |
|---|---|---|
| **The regulatory experiment is set up like other RD&D projects: It is in a specific area and carried out by specific stakeholders.** | | |
| **Option 1:** **Open call** | Call for tender for one (or more) experiments in a specific area, the definition of detailed regulatory question is part of the tender. | One advantage is that the potential regulatory question is generated directly from stakeholders during the bidding phase. Stakeholders are challenged with translating the issues into system-relevant questions, based on their respective business perspective. The regulator can then decide which proposal is best from its perspective. |
| **Option 2:** **Call for specific experiments** | Call for tender for one (or more) specific regulatory question based on a system perspective. | In a two-step process, relevant regulatory issues can first be identified together with stakeholders, and these issues can then be included in the call for tender. |
| **Option 3:** **Link to a demonstration programme** | The regulatory experiment takes place within a larger demonstration programme and is linked to the projects in this programme. | |
| **The regulatory experiment is carried out by the regulator itself.** | | |
| | There is no tender for a project in which the regulatory experiment takes place. The regulator itself oversees organizing and implementing the experiment. | The regulatory question should still be developed together with stakeholders, as well as the experiment itself. |

## 4.2.5   Preparing evaluation and learning

Experiments need to lead to learning and potentially upscaling the results, which requires an explicit evaluation process. **Evaluation** not only becomes relevant at the end of the project, but rather **needs to be set up at the very beginning**. This involves drawing up a concept for evaluation and providing the relevant resources so that the concept can be executed in the following phases. For example, a scientific and independent monitoring and evaluation of the experiment can increase learning. This process should not depend on government changes, for example.

It also implies that the experiment should be set up in line with the evaluation requirements, so that it becomes possible to achieve and evaluate the required results. This includes the following issues.

**Type of regulatory experiment**
These issues are mainly relevant for Regulatory Innovation Trials.

Formulate an explicit hypothesis that can be tested.

- Define useful indicators.
- Should only one regulatory option be tested, or several options in parallel so that the effects can be compared?

Experience shows that testing several design options in regulatory experiments increases learning.

**Case example**
The Chinese emission-trading pilot scheme can be seen as a best practice for testing several design options in parallel. The authorities simultaneously started seven pilot projects, each of which were given substantial leeway in the design of the regional emission-trading schemes. It shows how experimenting with a multitude of pilot projects provides more possibilities to learn during the experiment and constantly improve the instrument tested, and enables comparison between the parallel experiments.

In the basic income experiment in Finland, the regulator decided to only test one partial basic income model with a monthly income of 560€. The research group around the Finnish social insurance authority Kela had suggested testing several concepts of basic income. As a result, the instrument tested could only be compared with the status quo, rather than the effects of an alternative option.

(see Annex for list of cases)

- Randomization and control groups: For the evaluation of the experiment, randomization procedures and a control group can be important. For regulatory sandboxes, sample randomization and control group implementa-

How to design and evaluate a Regulatory Experiment?
A Guide for Public Administrations

**19**

tion are not important design elements. In such cases, a transparent selection process based on well-defined criteria seems to be important to achieve the goals of the experiment and ensure learning processes. It may suffice to define the control group ex-post (e.g. by comparing with actors who do not participate). Control groups are especially difficult to assign if participation is voluntary (see Chapter 4.2.3 How to involve stakeholders).

**Case example**

The Finnish basic income experiment deliberately used randomization procedures to form treatment and control groups. The experimenters used register data of the Finnish social insurance authority to randomly assign 2,000 individuals of the target group to the treatment group and the remaining 175,000 individuals to the control group. Participation in the experiment was mandatory. This design enabled gaining knowledge about the effects of basic income on several variables, e.g. employment.

By contrast, the basic income experiment in Ontario did not use randomization to form treatment and control groups. Individuals had to apply to participate in the experiment. Furthermore, the experiment did not cover whole of Ontario but three regions that are representative of the population. In two of the regions, 1,000 individuals were assigned to the treatment groups while the third region served as a saturation site, where approximately 85% of the target group received basic income.

One can conclude from this that the design of the Finnish basic income experiment allowed measuring causal effects of the tested model of basic income because individuals were randomly assigned to treatment and control groups, which was not the case in the experiment in Ontario. However, the implementation of a saturation site in Ontario would have allowed measuring community-level effects had the experiment not been cancelled.

(see Annex for list of cases)

The following figure summarizes key evaluation issues for the different phases that should be covered in the evaluation concept.

| **Goal and type** ▶ | **Preparation** ▶ | **Implementation** ▶ | **Evaluation** |
|---|---|---|---|
| <ul><li>Define the reference for evaluation</li><li>Societal and sustainability relevance?</li><li>Goal of the experiment?</li><li>Define clear regulatory question</li></ul> | <ul><li>Draw up evaluation concept</li><li>Embed experiments in broader learning context</li><li>Design experiment in line with evaluation requirements (e.g. control group, duration of the experiment)</li></ul> | <ul><li>Monitor the implementation</li><li>Be ready to fine-tune the experimental design</li><li>If needed, even adapt instrument design during the implementation</li></ul> | <ul><li>Evaluate the experiment to learn about the tested instrument and how it can be upscaled</li><li>Recommend regulatory option(s)</li><li>Adapt regulatory option based on evaluation results</li><li>Prepare the up-scaling of the results</li></ul> |

### 4.2.6 Size of the experiment

Experiments typically are small-scale trials, but regulatory experiments can also be large scale, and a new instrument can be tested in a country; for example, a new auction scheme for renewable energy systems. The size of the experiment is an important decision.

The size can refer to the geographical area in which the regulatory experiment takes place, such as a region. Nonetheless, the experiment may not be limited in size, but only in time.

The question of size is also linked to the experimental design; for example, whether there is only one project or several projects in parallel that can also learn from each other (see Chapter 4.2.5).

Finally, the number of participants also affects the size of the experiment (see Chapter 4.2.3 Stakeholder involvement).

### 4.2.7 Duration of the experiment

By definition, an experiment is limited in time. On the one hand, this causes the general problem that it can influence the effects the tested policy instrument has. On the other hand, clear timeframes can also foster and focus learning processes. Accordingly, how long should the experiment be? This should be carefully considered when setting up an experiment.

> **Type of regulatory experiment**
> The typical timeframe substantially differs between Regulatory Sandboxes and Regulatory Innovation Trials. Regulatory Sandboxes are embedded in permanent regulatory framework where fixed time frames exist for testing. Regulatory Innovation Trials are often planned in a predefined time frame. During the implementation, the duration of experiments can deviate from the planned timeframe due to changing context conditions.

If experiments are too short, it may be difficult for the experimenters to obtain useful results. It may not be possible to adequately evaluate the effects of a tested policy and how the policy can expect to work in a different or broader context. This can be especially true for more long-term effects. The artificial end of experiments can shape participants' behavior and thus bias the results, particularly with shorter experiments. More time is also needed if new insights are already to be taken into account during the experiment.

If the tested instrument especially addresses investment decisions, the fixed-term nature of the experiment is particularly problematic if the payback period for the investment exceeds the duration of the experiment. In this case, one option could be that only during the experiment can the new rules be applied for investments, but once the investment is made the rule covers the whole payback period and not only the duration of the experiment.

Nonetheless, if experiments are too long, this can also cause problems. Regulatory experiments – as real-world experiments – cannot be conducted in complete isolation from broader political and societal changes. In order to minimize the influence of such factors that hamper a causal interpretation, it is necessary to limit the duration of regulatory experiments. In addition, rising costs due to the long duration can become a barrier for conducting the regulatory experiment. Long regulatory experiments also entail the risk of discontinuation for political reasons. **Regulatory experiments that are too long may also not generate results on time to feed into the further political process and may even delay the introduction of new solutions.**

> **Case example**
> The analysed cases lasted from one to ten years, with the majority lasting up to three years.
>
> The pilot tenders for the promotion of renewable energy in Germany and the basic income experiment in Ontario have shown that the effects of a tested policy cannot be adequately evaluated if the period of an experiment is too short. In the former case, the evaluator was unable to assess the realization rate of successful bids. In the latter case, the experiment was canceled prematurely due to a change in government.
>
> Experimenters also have to consider how the political election cycle influences the setup of the experiment. For example, the Finnish Prime Minister ordered experimenters to conduct the basic income experiment earlier than planned so that the results could be presented ahead of the elections. With the same intention, the timeframe of the Ontario basic income experiment was set shorter than what may have been appropriate.
>
> **(see Annex for list of cases)**

When deciding on the duration of the experiment, there should be sufficient time for preparing the experiment, e.g. getting stakeholders on board, as well as evaluating the experiment.

## 4.3 Phase 3: Implementation

### Implementation

- ➡ Monitor the implementation
- ➡ Be ready to fine-tune the experimental design
- ➡ If needed, even adapt instrument design during the implementation
- ➡ Manage stakeholder involvement

The implementation phase depends on the design that is defined in the previous phase. Nonetheless, it should not be a mere execution of a predefined programme. Especially as experiments aim at learning, there cannot only be an experimental design that is set up ex-ante; rather, learning should be an ongoing process that can also lead to adapting the design or even the regulatory question or the tested policy instrument. This is especially true in the case of RITs. It is important to have a clear plan and a stable experimental setting, but one should also be able to adapt them if necessary.

This is especially true the higher the uncertainty concerning the new regulatory instrument. While in some cases it may be very clear which instrument is to be tested, in other cases the regulatory problem can be less well-known upfront and may only completely emerge during the experiment. In such cases, **flexibility helps to maximize the benefits of the experiments**. In several examples of regulatory experiments, monitoring mechanisms to react to developments during the experiment were put in place, in addition to ex-post evaluations. There can also be an explicit mid-term evalua-

tion. In some cases, monitoring can also lead to the termination of the experiment.

### Case example

**In addition to ex-post evaluations, some experimenters put in place monitoring mechanisms to react to developments during the experiment. For example, the energy sandbox in Singapore requires applicants to report to the regulator on the progress of the experiment. If the regulator loses confidence that the tested innovation fulfills its purpose or detects a substantial deficiency that cannot be corrected, the experiment can be terminated. This also applies to cases where the applicant breached conditions of the sandbox.**

**(see Annex for list of cases)**

Another important task during the implementation is to keep stakeholders on board, including dealing with potential resistance, benefiting from their input, engaging them in the experiment and fostering their learning.



photo: istockphoto / Lucia Gajdosikova

## 4.4  Phase 4: Evaluation and upscaling

**Evaluation and Upscaling**

➧ Evaluate the experiment to learn about
  ➧ the tested regulatory innovation
  ➧ how it can be upscaled
  ➧ the experimental design
➧ Recommend regulatory option(s)
➧ Adapt regulatory option based on evaluation results
➧ Prepare the upscaling of results

In the evaluation phase, the summary and analysis of the experimental procedure and outcomes are carried out. The evaluation phase is key to realizing regulatory learning in regulatory experiments. It is based on the planning of the evaluation described in phase 2.

Ideally, learning should happen on two levels: 1) learning about the tested regulatory innovation, its practical implementation and effects; and 2) learning about regulatory experiments so that this tool can be further developed, including in other contexts. Therefore, the evaluation should cover both the policy instrument itself as well as the experimental design. The latter is typically part of the program or project evaluation.

**Learning can refer to learning about the policy instrument from a scientific perspective, it can also mean learning for different types of stakeholders – for example, regulatory or entrepreneurial learning – and it can refer to interactive learning that affects actors' behavior regarding information acquisition, communication and cooperation.**

In this phase, evaluation results from the experiment need to be combined with the results generated with other tools such as role plays or modeling, as well as other new insights that have been gained while the experiment has been running.

**The evaluation needs to consider the effects that can be observed when new regulation is implemented in practice, how its implementation can be improved, which un-** **intended consequences can be observed and how these can be avoided.** The monitoring approaches applied in the experiment can also help design the monitoring mechanisms that should be included in any new legislation.

If possible, the evaluation should not only cover the effects of the tested regulatory option(s) that could be observed during experiment, but it should also draw conclusions on how to refine the instrument before upscaling it or applying it in different contexts. If several alternative regulatory options have been tested in parallel, the evaluation should recommend one of them or explain which option fits best under which conditions.

The evaluation should also look beyond the experiment. Based on the results of the experiment, how is the policy instrument expected to work when it is upscaled – for example, to the national level? Under which conditions or in which scenarios can it be transferred (including to other jurisdictions on the same governance level)? How does the instrument need to be adapted for this purpose? On top of recommendations concerning the design of the instrument, the evaluation should also address the process of upscaling. **Which steps can be taken to introduce the instrument on a larger scale?** As experiments relate to politics and the expectations and interests of stakeholders involved, the evaluation should also cover this dimension and draw conclusions for the implementation of the tested regulatory innovation.

# Annex:
# List of project publications and case studies

## Publication list

Bauknecht, D., Bischoff, T. S., Bizer, K., Führ, M., Gailhofer, P., Heyen, D. A., Proeger, T. & Von der Leyen, K. (2020). **Exploring the pathways: Regulatory experiments for Sustainable Development – an interdisciplinary approach.** Journal of Governance and Regulation 9 (3), 49-71. **Download**

(https://www.virtusinterpress.org/Exploring-the-pathways-Regulatory-experiments-for-sustainable-development-An.html)

Bischoff, T. S., Von der Leyen, K., Winkler-Portmann, S., Bauknecht, D., Bizer, K., Englert, M., Führ, M., Heyen, D. A., Gailhofer, P., Proeger, T. & Vogel, M. (2020). **Regulatory experimentation as a tool to generate learning processes and govern innovation – An analysis of 26 international cases.** sofia-Diskussionsbeiträge, 20-7, Darmstadt. **Download**

(https://www.sofia-darmstadt.de/fileadmin/Dokumente/Diskussion/2020/Netzversion__Portmann_Regulatory.pdf)

Feser, D., Winkler-Portmann, S., Bischoff, T. S., Bauknecht, D., Bizer, K., Führ, M., Heyen, D. A., Proeger, T., von der Leyen, K. & Vogel, M. (2021). **Institutional conditions for the up-take of governance experiments – A comparative case study.** ifh Working Papers (No. 28). Göttingen. **Download**

(https://www.ifh.wiwi.uni-goettingen.de/upload/veroeffentlichungen/WP/ifh_wp-28_2021.pdf)

## Case studies

| (a) Regulatory Sandboxes | | (b) Regulatory Innovation Trials | |
|---|---|---|---|
| 1 | Austria: Regulatory sandbox in the electricity sector (Energy.Free.Room) | 13 | China: Trading scheme for $CO_2$ emissions |
| 2 | Australia: Regulatory sandbox in the electricity sector | 14 | Germany: Tender procedure for promoting renewable energy systems |
| 3 | Singapore: Regulatory sandbox in the electricity & gas sector | 15 | Pennsylvania, US: Promoting grocery stores in under-served neighborhoods (Fresh Food Financing Initiative) |
| 4 | Germany: Smart meter standardization | 16 | India: Immunization access and incentives |
| 5 | Netherlands: Regulatory sandbox in the electricity sector | 17 | Norway: Promoting electro-mobility |
| 6 | United Kingdom: Regulatory sandbox in the electricity & gas sector UK | 18 | Berlin, Germany: Shared space pilot projects (Begegnungszonen) |
| 7 | Norway: Exemption clauses in the energy regulation | 19 | Copenhagen, Denmark: Promoting biking |
| 8 | Germany: Retroactive reimbursement in the electricity sector (SINTEG Ordinance) | 20 | Germany: Occupational licensing in the crafts sector |
| 9 | Denmark: Regulatory sandbox in the financial sector | 21 | Indonesia: Labor market consequences of school construction |
| 10 | United Kingdom: Regulatory sandbox in the financial sector | 22 | Finland: Basic income experiment |
| 11 | European Union: Authorization scheme for the use of chemicals | 23 | Ontario, Canada: Basic income experiment |
| 12 | Italy: Derogations and pilot projects in the electricity sector | 24 | Seattle, United States: Minimum wage policy |
| | | 25 | Ontario, Canada: Minimum wage policy |
| | | 26 | Berlin, Germany: Solidary basic income experiment |
| | | 27 | Barcelona, Spain: Decentralized citizen-owned data ecosystem (DECODE) |

AUTONOMOUS DRIVING

**How to design and evaluate
a Regulatory Experiment?**

A Guide for Public Administrations