



Volume 1, Number 2 - Fall 2015

## Table of Contents

<b>Editorial Introduction to Second Issue of <i>European Policy Analysis</i> (EPA)</b> _____	<b>1</b>
<b>Risk-based Regulation in Continental Europe? Explaining the Corporatist Turn to Risk in German Work Safety Policies</b> _____	<b>5</b>
<i>Regine Paul, Law and Society Unit, University of Bielefeld, Germany</i> <i>Michael Huber, Law and Society Unit, University of Bielefeld, Germany</i>	
<b>Crisis, Policy Discourse, and Major Policy Change: Exploring the Role of Subsystem Polarization in Nuclear Energy Policymaking</b> _____	<b>34</b>
<i>Adrian Rinscheid, University of St. Gallen, Switzerland</i>	
<b>Regulation and the Management of Expectations – Rating Agencies Revisited</b> _____	<b>71</b>
<i>Markus Hinterleitner, Center of Competence for Public Management, University of Bern, Switzerland</i> <i>Christian Rosser, University of Konstanz, Germany &amp; University of Lausanne, Switzerland</i>	
<b>Addressing Multilevel Program Complexity by Evaluation Design</b> _____	<b>90</b>
<i>Fritz Sager, Center of Competence for Public Management, University of Bern, Switzerland</i> <i>Céline Mavrot, Center of Competence for Public Management, University of Bern, Switzerland</i> <i>Susanne Hadorn, Center of Competence for Public Management, University of Bern, Switzerland</i>	
<b>The Use of Evidence Within Policy Evaluation in Health in Ghana: Implications for Accountability and Democratic Governance</b> _____	<b>111</b>
<i>Elisa Vecchione, London School of Hygiene and Tropical Medicine (LSHTM), United Kingdom</i> <i>Justin Parkhurst, London School of Hygiene and Tropical Medicine (LSHTM), United Kingdom</i>	
<b>Law Activation Strategies (LAS) in Environmental Policymaking: a Social Mechanism for Re-politicization?</b> _____	<b>132</b>
<i>Rémi Schweizer, Institute of Political, Historical and International Studies (IEPHI), University of Lausanne, Switzerland</i>	
<b>The Communication Behavior of German MPs on Twitter: Preaching to the Converted and Attacking Opponents</b> _____	<b>155</b>
<i>Simon Hegelich, ForschungsKollegSiegen (FoKoS), University of Siegen, Germany</i> <i>Morteza Shahrezaye, ForschungsKollegSiegen (FoKoS), University of Siegen, Germany</i>	



# Addressing Multilevel Program Complexity by Evaluation Design

Fritz Sager,<sup>A</sup> Céline Mavrot<sup>B</sup> & Susanne Hadorn<sup>C</sup>

*In this paper, we present the evaluation design for a complex multilevel program recently introduced in Switzerland. The evaluation embraces the federal level, the cantonal program level, and the project level where target groups are directly addressed. We employ Pawson and Tilley's realist evaluation approach, in order to do justice to the varying context factors that impact the cantonal programs leading to varying effectiveness of the implemented activities. The application of the model to the canton of Uri shows that the numerous vertical and horizontal relations play a crucial role for the program's effectiveness. As a general learning for the evaluation of complex programs, we state that there is a need to consider all affected levels of a program and that no monocausal effects can be singled out in programs where multiple interventions address the same problem. Moreover, considering all affected levels of a program can mean going beyond the borders of the actual program organization and including factors that do not directly interfere with the policy delivery as such. In particular, we found that the relationship between the cantonal and the federal level was a crucial organizational factor influencing the effectiveness of the cantonal program.*

**Keywords:** *evaluation design, multi-level governance, complexity, realist evaluation, implementation analysis, tobacco prevention program*

## Introduction

The present paper presents the approach used in a comparative evaluation of the new cantonal tobacco prevention programs in Switzerland and links this case study with current trends in evaluation practices. The complex frame of the Swiss tobacco prevention program offers an interesting case for various dimensions of policy evaluation. Within a

set of funding criteria defined at the federal level, sub-national administrations are incentivized to frame their own, customized prevention programs. The main expectation of the involved tobacco prevention experts about the policy is that coordinated programs lead to synergies and hence better outcomes than uncoordinated projects. As a result of this allocation of responsibilities, each cantonal program consists of an individually composed set

---

<sup>A</sup> Center of Competence for Public Management, University of Bern, Switzerland

<sup>B</sup> Center of Competence for Public Management, University of Bern, Switzerland

<sup>C</sup> Center of Competence for Public Management, University of Bern, Switzerland

of projects with a distinct emphasis on structural prevention, behavior oriented prevention, and information activities. An important question of the evaluation is whether the cantons are the appropriate level of action to adopt effective policies.

The evaluation concept of these cantonal programs is discussed in this paper. The aim is to contribute to the advancement of complex program evaluation designs. We claim to achieve this, as the evaluation has to focus on different levels in the case at hand. Firstly, the individual projects have to be assessed; secondly, each cantonal program will be subject to evaluation. Lastly, in a future step, 14 program evaluations will be used as a basis for an inter-cantonal comparison, where the composition of the programs as well as the impact of the specific cantonal context will be studied more closely. The upcoming comparative analysis will thus strive at presenting explanations for cantonal differences in efficacy and efficiency, both based on the policy concept as well as the implementation phase. In the present article, we present our evaluation concept as well as a first case study.

The paper is structured as follows: it starts with a short introduction to decentralized policy implementation in Swiss federalism in which the cantonal tobacco prevention programs are situated. We then present our evaluation design with which these cantonal programs will be assessed and illustrate its realization with the findings from the evaluation of the canton of Uri. Ultimately, the paper endeavors to provide some conclusions about the effects of policy co-formulation as well as the challenges of multi-level governance with regard to the selected case study.

## **Swiss federalism and implementation by federal delegation**

**I**n Switzerland, federalism is the most distinctive feature of the political system after direct democracy (Sager and Zollinger 2011). Only a few West European countries have systems in which not only the federal government but also the constituent state enjoys legislative, executive, and judicative power (Ismayr 2009). In addition to Switzerland, traditionally, this can be observed in Germany and Austria and more recently in Belgium and Spain.

The implementation of federal policies in Switzerland is strongly shaped by this institutional setting. The federal programs are implemented by the cantons and communes. “While the Federation holds the legislative power in many areas, responsibility for implementing federal policies resides to a large extent with the cantons” (Vatter 2007, 91). Therefore, in a great number of policy domains, the federal level is dependent on the cantons for the implementation of federal legislation. Due to the high degree of legislative autonomy of the Swiss cantons, the delegation applies not only to the actual implementation of federal laws (i.e., the right to act), but also to the adaption of these provisions to the local situation (i.e., the right to decide). The cantons are not only implementing, but also programming authorities.

According to Linder (1987; 2010) the implementation of federal policies by the cantons is advantageous both for the federation and for the cantons. Whereas for the central government the advantage lies in reducing its workload, for the cantons, the advantage lies in being able to

control their own program priorities and in adapting policy implementation to the local context. Despite this optimistic view, “implementation by federal delegation” draws mixed reviews (Kissling-Näf and Wälti 2007, 504). While some lament the lack of federal control over implementation, others value the receptivity and adaptability as well as the flexibility and experimental character of decentralized implementation. A further drawback is that not every canton can afford to provide for well-funded and professionalized implementation units (Sager 2003). Moreover, authors have shown the crucial importance of the issue framing and of the use of evidence during the different local policymaking processes, sometimes leading to very distinct cantonal outcomes (Balthasar and Müller 2014; Blatter, Bombach, and Wiprächtiger 2015). The central role of the horizontal cooperation among federated entities for the adoption of fragmented or countrywide policies in federalist countries has also been evidenced (Rothmayr, Varone, and Montpetit 2003). Consequently, an additional problem that derives from the “implementation by federal delegation” is uneven results between the cantons (Sager and Rüefli 2005).

Both advantages and drawbacks also come into play when policy programs launched at the federal level are meant to lead to coordinated policy action at the cantonal level. For the case of palliative care, Moser and Sager (2015) show cantonal path dependencies to play a crucial role for whether cantons respond to vertical stimulus or not. As regards drug policies, Kübler (2000) demonstrates the importance of the local advocacy coalitions, and Mavrot (2012) highlighted the role of the political and health professional constellations in modeling the canton-

al policymaking process. The drug policy example also reminds that the existence of a relatively independent metropolitan governance level should neither be neglected (Kübler and Wälti 2001). Finally, in the case of alcohol control policy, Sager and Rielle (2013) identify administrative structures as a core factor for the adoption of cantonal policy programs. Accordingly, Mavrot and Sager (2014) argue so-called vertical epistemic communities (national and cantonal experts) to play a major role for the establishment of cantonal tobacco prevention programs. The case of these programs is presented in the following section.

### **The Tobacco Prevention Fund and the cantonal tobacco prevention programs**

**T**obacco prevention policy in Switzerland took only off after 2001 with the National Tobacco Prevention Program 2001–2005. However, until the early 2000, due to the central state weakness regarding policy implementation, coordination at the cantonal level mainly derived from voluntary inter-cantonal efforts. Nevertheless, considerable change has taken place since the creation of the Tobacco Prevention Fund in 2004 as a leading national actor. Loose and partial inter-cantonal coordination is gradually being replaced by a more vertical coordinated approach including all aspects of smoking prevention.

The Fund was created after the national parliament decided to revise the tobacco taxation legislation in 2003. Taxes increased, whereby an important source of income dedicated to smoking prevention was obtained. The Fund was created to

manage this new income by funding any relevant smoking prevention project both those from nongovernmental organizations (NGOs) and those from cantonal administrations. This aspect of its mission gives the Fund a significant weight in influencing the policymaking. Its decision-making power in financial allocation allows the Fund to act proactively. The strategic orientation of the Fund is provided by the National Tobacco Program, which is designed by a broad panel of actors under the leadership of the Federal Office of Public Health (FOPH) and approved by the federal government on a regular basis. It includes the following lines of action: preventing the beginning and promoting the cessation of consumption, protecting against second-hand smoke exposure, raising public awareness, encouraging research activities, developing framework conditions to foster preventive actions, and finally, building a national network of tobacco prevention actors (RS 641.316, art.2). The Fund managed to gain great influence in the relevant policymaking processes despite its proportionally weak resources in terms of personnel. Since its creation, its collaborator numbers have always been very low for an organization deciding the allocation of millions of francs. Hence, the two major strengths of the Fund are: (a) its significant budget and its privilege to choose which prevention projects will be funded; and (b) the large network of experts it has rapidly built. The Fund's high financial capacity and authority to make allocation decisions provide it with considerable power of persuasion and of steering (Mavrot and Sager 2014).

The Fund decided to promote two aspects: reinforcing the role of the cantons in order to foster centrally coordinated public policies rather than scattered activ-

ities, and enhancing the evidence-based nature of tobacco prevention in Switzerland. The resulting new scheme for cantonal tobacco programs is that cantons can apply for considerable financial resources from the Fund when proposing an overall evidence-based set of prevention projects that will be coordinated by a cantonal program office and implemented by various actors from the tobacco prevention policy network. The projects can be either managed by public or private entities. The Fund expects the program format to result in more coordinated action and synergy effects that will lead to better effectiveness in terms of outcomes and impact. At the same time, we can observe that this scheme leads to very complex horizontal and vertical relations that impact program effectiveness (see Figure 1). Firstly, the Fund at the federal level horizontally needs to coordinate with the FOPH that demands that cantonal programs are in line with the national tobacco prevention policy. Secondly, the Fund is aiming at coordinating the various cantonal efforts by the selective allocation of funds. However, not only cantons can apply for project funding but also NGOs that act at the national scale. In comparison to the cantons, NGOs are the incumbent players of tobacco prevention in Switzerland. In the case where such NGO projects are implemented in a canton applying for a program, these national projects also need to be coordinated even though they were originally not designed to be part of a cantonal program.

Thirdly, cantons question the Fund's legitimacy to steer their programs as the tax money does not belong to the federal level but is simply bound to be spent on tobacco prevention. Consequently, the Association of Cantonal Pre-

National Tobacco Prevention Program

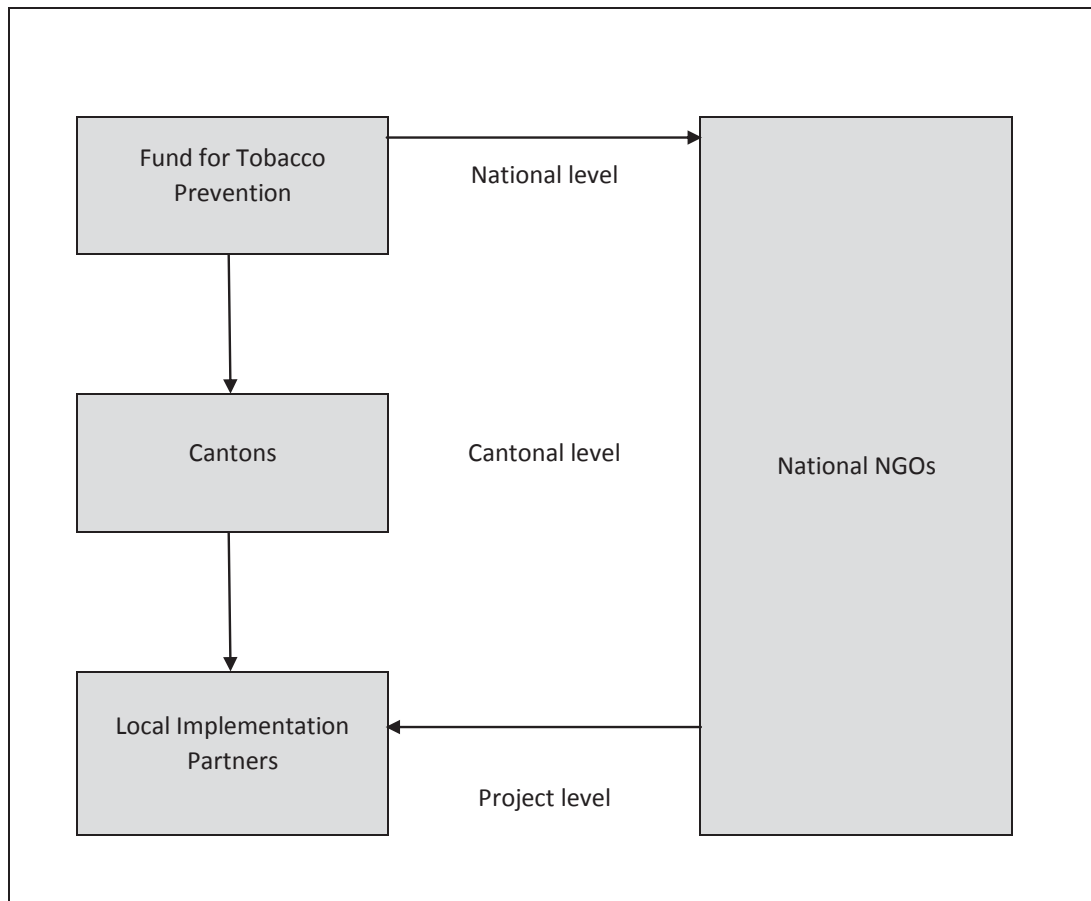


Figure Legend:

→ = financial flow

Figure 1: Swiss Tobacco Prevention System



ventions Experts<sup>1</sup> acts as sounding board for the cantonal programs and interferes with the Fund's policy when deemed necessary. Fourthly, which is also new, cantons implementing a program have to internally coordinate with all actors engaged in tobacco prevention activities in the canton (local NGOs, all involved administrative units, etc.). The cantonal coordinators running the programs are in a difficult double role in which they have to manage the variety of projects in their canton and at the same time are accountable to the Fund. This double role is even more difficult as most of the cantonal coordinators are not managers but prevention specialists, i.e., neither trained in managing large projects nor in monitoring and collecting evaluation data they are bound to provide to the Fund in order to get the funds. Fifthly, the Swiss cantons are fully fledged political systems that vary not only in social, economic, and cultural aspects but also in terms of politics and administrative organization. Context hence plays a crucial role for program effectiveness.

### **Capturing vertical complexity: a realistic multilevel evaluation design**

**T**he complex frame of the Swiss tobacco prevention program offers an interesting case for various dimensions of policy evaluation. It is, for example, necessary to examine if the sub-national administrations, which are now entitled to frame their own tailored program, are the appropriate level of action to design and implement effective smoking

prevention policies. Each cantonal program includes a range of preventive projects, as well as monitoring and evaluating components. The cantonal programs are required to achieve a balance between the different types of prevention (structural prevention, behavior-oriented prevention, and information). The same applies to the balance between the prevention of the beginning of use, the promotion of smoking cessation, and the protection against second-hand smoke. Finally, the projects have to cover a variety of settings such as for example schools, working environment, sports, or the medical field. The aim is to develop the most appropriate array of activities, based on the cantonal realities.

Given the depicted framework, the qualitative evaluation focuses on different levels. Firstly, each individual project within the programs is assessed. Secondly, each cantonal program will be evaluated. Both the policy concept and the policy implementation will be taken into account, as we will illustrate here with the example of the canton of Uri. In the future, we will analyze 14 programs within the framework of an inter-cantonal comparison. In this upcoming study, particular attention will be paid to the composition of the different programs, as well as to the varying cantonal contexts impacting their setting up. Thus, this comparative analysis will aim at understanding the cantonal differences regarding the overall efficacy and the efficiency of the smoking prevention programs. Consequently, the whole evaluation of the cantonal programs is not mandated

---

<sup>1</sup> Vereinigung der kantonalen Beauftragten für Gesundheitsförderung in der Schweiz – VBGF.

in the form of an evaluation contract but financed by the Fund as a basic research project.

The main questions with regard to the single cantonal programs are

- Which, if any progress, can be stated in tobacco prevention (i.e., intended changes within target groups, reduced prevalence)?
- Has the program met its goals (i.e., prevention and reduction of consumption, information of the public, protection against second-hand smoke)?
- Which are the drivers and which are the obstacles to progress in tobacco prevention?
- Which, if any, is the added value of the coordinated program as opposed to previously uncoordinated projects?
- Overall, was the cantonal program a success (i.e., led to the desired outcomes and impacts)?

In the following, we first present our general theoretical approach to this evaluation, namely realistic evaluation, before we turn to our analytical model, the different evaluation elements and the respective hypotheses.

#### *General approach: realistic evaluation*

Realistic evaluation stresses the importance of the context in the analysis of public policies (Pawson and Tilley 1997). Realistic evaluations decompose political programs into three components: the program context (C), mechanism (M), and outcome (O). According to this approach, political programs can generate a whole range of different mechanisms of

change. The specific mechanisms triggered depend on the program context. This notion of complex causation means that the relationship between causal mechanisms and their effects is not predetermined, but contingent. Context matters because it “turns (or fails to turn) causal potential into causal outcome” (Pawson and Tilley 1997, 69). As a result, a political program can lead to a variety of different outcomes according to the context. This translates into different context–mechanism–outcome configurations, referred to as CMO configurations. Realistic evaluation conceptualizes and tests these CMO configurations in order to understand what works for whom and under what circumstances (Sager 2008).

Pawson and Tilley (1997, 159) underline the importance of building on theory in defining the focus of an evaluation. Moreover, the use of theory and of the results of previous studies in a given field fosters knowledge accumulation. Pawson (2002a; 2002b) further transfers this logic of theory-driven and context-sensitive evaluation to the level of research synthesis. According to him, the accumulation of knowledge could be enhanced through the “realist synthesis”: a summary of previous research that follows the CMO pattern, sorted according to the mechanisms they detected, the contexts they analyzed, and the outcomes they were able to observe. The result would be a gathering of similar CMO configurations which could be merged into more abstract ones. Pawson’s conception of context relies heavily on the idea of contingency. This idea implies that the meaning of social action depends on the specific conditions in which it occurs. Accordingly, the perception of the addressees of policy interventions defines the meaning of social action and



thus determines the outcome of a mechanism within a CMO configuration (Befani, Ledermann, and Sager 2007; Sager 2007; Sager and Andereggen 2012). In the following, we present the different elements of our evaluation model and formulate the related hypotheses.

*Evaluation model*

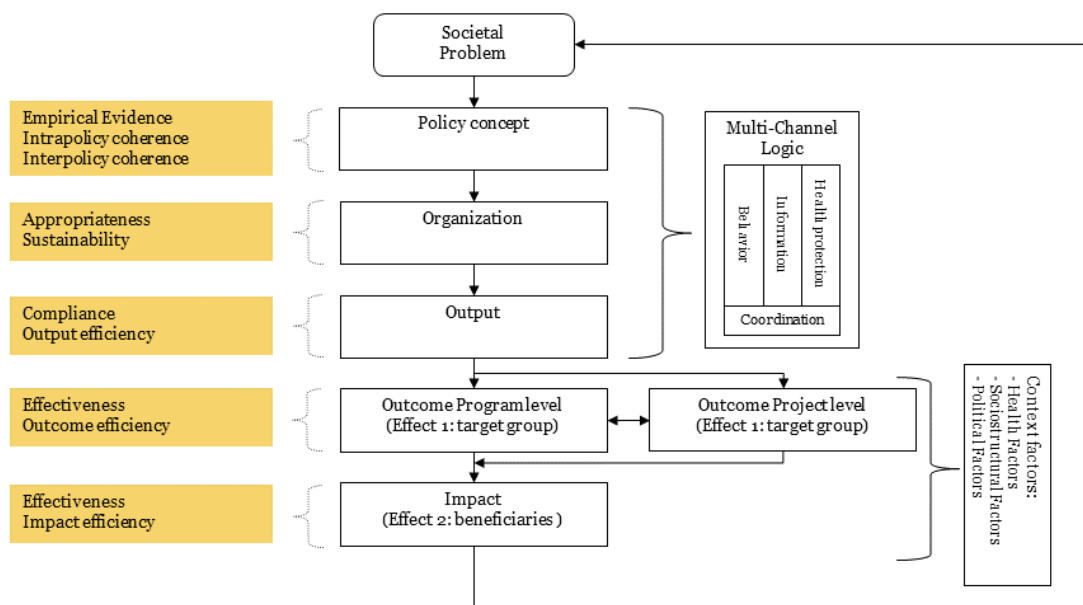
Our general model for cantonal programs is presented in Figure 2. The focus of this evaluation is on the program level and we therefore refrain from hypotheses at the project level. However, the actual impact of the project only takes place at this latter level where the final target groups are addressed. We therefore use

project data to measure outcome and impact.

In what follows, we define the different analytical stages, name the employed evaluation criteria and formulate causal hypotheses.

*Policy concept*

The policy concept embraces the causal model of a program and the planned implementation structure as put to paper in the cantonal proposals. The concept hence is a program theory before it is put into practice (Rossi, Freeman, and Lipsey 1999, 156; Sager and Hinterleitner 2014, 444). The criteria we use to evaluate the concept are fourfold: empirical evi-



Adapted from Sager and Hinterleitner (2014, 444)

Figure 2: Evaluation model

dence, precision, internal coherence, and external coherence (Knoepfel et al. 1997; Ledermann and Sager 2009). Empirical evidence refers to the backing of the assumed effects on scientific studies in the fields of public health research and policy analysis, or on policy experience previously gained by the local actors (cf. Howlett 2009). Precision concerns the accuracy of the program objectives all the way to the final intended effects. Internal coherence means the fit of the different elements of a program, i.e., the avoidance of internal contradictions in the program logic and potential implementation deficits (Howlett and Rayner 2007). External coherence refers to the policy context of a program and asks whether the program is in line with other relevant programs in the canton that might interfere with the program logic (May, Sapotichne, and Workman 2006). This regards both other prevention programs and programs from other policy areas such as economic or education policy. An appropriate policy concept is also expected to account for the distinction between policy failure and implementation failure in that organizational aspects should be planned in accordance to the expected causal effects of the foreseen interventions (Ledermann and Sager 2009). The hypotheses for the policy concept focus on the effectiveness of the program:

1. The more a program is based on empirical evidence, the better its effectiveness with regard to (a) outcome and (b) impact.
2. The more precise the objectives of a program are defined, the better its effectiveness with regard to (a) outcome and (b) impact.
3. The higher the internal coherence of a program, the better its effective-

ness with regard to (a) outcome and (b) impact.

4. The higher the external coherence of a program, the better its effectiveness with regard to (a) outcome and (b) impact.

#### *Organization*

Organization refers to a cantonal program's implementation structure as in fact established during the implementation. The realized implementation structure of a program does not necessarily correspond to the planned implementation structure. We understand organization as an institutional structure (Egeberg 2003). Institutions distribute decision power by rules and routines and coin actor identities and their interpretations of situations. They are therefore crucial for policy delivery (Nicholson-Crotty and O'Toole 2004). Organization entails the allocation of decision and action competencies, financial and personnel means, and expertise (Sager 2004; Sager and Rielle 2013). Evaluation criteria are appropriateness and sustainability of an implementation structure. Appropriateness refers to the question of how far the organization is in line with the actual tasks for service provision. Particularly important in this respect is the coordination capacity. Sustainability regards long-term security of the programs and projects in order to avoid stop-and-go problems damaging effectiveness. The hypotheses read as follows:

5. The more appropriate the implementation structure of a program, the better its outcomes.
6. The more sustainable the implementation structure of a program, the better its outcomes.

We now turn to the output side of the program. Following Patton (1997, 193–194), the success of public policy, thus understood, can be measured on the three levels: outputs, outcome, and impact.

### *Outputs*

Outputs represent the products of a policy. These are public interventions or services that attempt to change actors' behavior (Patton 1997, 193; Sager and Hinterleitner 2014, 446). To achieve this, outputs create a direct relationship between the competent implementation actor and the political target group, where the political target group refers to the group of actors whose behavior is seen by public politics as relevant to the problem in question (Sager 2007). If the planned outputs are not delivered in the intended manner in terms of time or scope, the policy cannot be expected to achieve its results due to implementation failure. At the program level, outputs basically regard management services toward the project implementers, i.e., governance and performance control, contract management, internal and external information and coordination.

Evaluation criteria for output performance are compliance and efficiency. However, as the latter does not have a direct relation to effectiveness but only considers the cost–service ratio, we do not use the efficiency criterion in this evaluation. Compliance refers to the degree to which planned services are actually delivered. This results in the following hypothesis:

7. The more the program implementation complies with the program concept, the better its outcome.

### *Outcome and impact*

Outcome designates the changes in the behavior of the political target group brought about by the outputs (Patton 1997; Sager and Hinterleitner 2014, 447). The addressees of a given policy are not necessarily equal to the beneficiaries of the policy. Rather, they are the ones the policy sees as responsible for causing the problem at stake or as able to help in addressing it. It is thus the addressees' behavior that can or needs to be changed in order for the groups suffering from a given problem to benefit from the change. For example, in order to protect small children in traffic the attention of drivers has to be increased. The target group's reactions to output may be as expected, take on a non-intended form, or indeed not take place at all.

An important issue in this respect with regard to prevention policies is the so-called “multi-channel” logic. The “multi-channel” logic states that behavior change in target groups is the result of a combination of multiple channels such as regulation, incentives, and information that may be limited to tobacco prevention or may go beyond substances and focuses on addictive behavior and its conditions as such (Bala et al. 2013; Lund and Aarø 2004). Hence, a prevention program is expected to reach the target groups by combining different fields of action operating in a variety of settings.

Finally, impact implies the totality of intended or non-intended effects which the achieved changes in the policy's target group's behavior have on the problem which the relevant policy is aiming to solve (Patton 1997; Sager and Hinterleitner 2014, 447). If the addressees change their behavior in the intended way with-

out an effect on the problem situation, the causal hypotheses of the policy is likely to be wrong (policy failure). Where addressees' behavioral change has a positive effect on the problem and an effect that corresponds to the political objective, one refers to an effective policy.

The criterion is effectiveness, i.e., the degree to which expected effects in term of outcome or impact are in fact achieved by the delivered outputs. Again, efficiency shall not be considered. The hypotheses refer to the causal link between the two effect stages, as well as to the "multi-channel" logic:

8. The higher a program's outcome effectiveness, the higher its impact effectiveness.
9. The stronger the "multi-channel" logic of a program, the higher its impact effectiveness.

### Context

Policy programs always take place in a specific context (Pawson 2002a; 2002b; Pawson and Tilley 1997). This context is defined by factors that exist exogenously and cannot be altered (or only with great difficulty), and which, to a certain degree, define the structure within which the program planning and implementation process takes place. The purpose of the inclusion of these contextual factors is such that, at the end, the success factors can be differentiated according to the framework of the project or of the program (Befani, Ledermann, and Sager 2007; Pawson and Tilley 1997; Sager and Andereggen 2012). This means that contextual factors impact the way target groups react to an intervention. Context hence is responsible for the fact that the same intervention once works

and another time does not. We therefore integrate context into our model and formulate three distinct general hypotheses regarding problem prevalence, socio-economic context, and political context.

10. Comparatively, large tobacco-related and general health problems in a canton increase the program's impact effectiveness.
11. An unfavorable socio-structural context in a canton decreases the program's impact effectiveness.
12. A favorable political context in a canton increases the program's impact effectiveness.

### Applying the model: findings for the canton of Uri

Within the scope of our mandate, Uri was the first canton subject to evaluation (Sager et al. 2015). The previously discussed evaluation model was applied to the case and the five elements: (1) policy concept, (2) organization, (3) output, (4) outcome and impact, as well as (5) context were analyzed. The main findings of the evaluation will be discussed below with reference to the defined hypotheses. In a nutshell, the program of Uri took the following form: it was made of 14 prevention projects, of which 10 were dedicated to youth, 2 to the general population and 2 to smokers. These projects were implemented in five different settings, in particular school and sport/leisure. They principally dealt with behavioral prevention, but covered structural prevention and information as well. Activities of all the implementation partners were centrally coordinated by the cantonal public health administration.

*Method*

The evaluation of the cantonal tobacco prevention program in Uri was based on the analysis of primary and secondary data, whereby the data sources differed between the five evaluation elements. More specifically, the evaluation element context was analyzed based on data provided by the Swiss Federal Statistical Office, national and cantonal health surveys, newspaper articles, economic reports, parliamentary debates as well as diverse cantonal documents. For instance, as a specific part of the context analysis, we conducted a media analysis including all articles published in the period between January 01, 2009 and December 31, 2014, in the “*Neue Urner Zeitung*,” using search terms relevant to tobacco prevention (such as “*Tabak*,” “*Passivrauch*,” etc.). Similarly, all relevant parliamentary debates at the cantonal level between 2009 and 2014 were included. The thereby collected documentation was assessed by means of a qualitative content analysis (Sager and Mavrot 2015). The analysis of the policy concept was mainly based on a content analysis of the canton’s prevention program proposal approved by the Fund. Last, the elements organization, output, outcome, and impact were assessed through cantonal documents, a time series of self-evaluation data provided by the cantonal implementing partners on a yearly basis during three years, as well as through additionally conducted semi-structured interviews with key players (e.g., cantonal program leader, national and cantonal project leaders, and local and sectoral implementation partners).

*Policy concept*

The case study of Uri fully confirmed the crucial importance of the policy concept underlying a program. This is the case regarding all four dimensions of the policy concept: the empirical evidence, the precision of objectives, as well as the internal and external coherence of the program (hypotheses 1–4).

The two main learnings regarding the empirical evidence (hypothesis 1) are the following: firstly, the canton based the concept more on its own past experience than on the available scientific evidence. The use of its own experience proved to be a positive factor. It allowed building on existing activities, to avoid mistakes made in the past, and to better identify the prevention gaps in the canton. Making an initial assessment of the cantonal situation with all implementation partners was of particular importance in this context where cantons are seeking for a better coordination of activities through “cantonalized” prevention programs. It proved to be a good opportunity for promoting concerted efforts. However, the use of scientific evidence regarding tobacco prevention was lower. Secondly, it turned out that when evidence was used—whether scientific or stemming from experience—this was more at a macro program level than for the design of specific projects. This lack of consideration of evidence at the project level led to some policy failures reducing the effectiveness of the program.

The precise definition of the objectives (hypothesis 2) also appeared to be a crucial part of a sound policy concept. Interestingly, the canton showed more ability to formulate precise objectives at the micro (i.e., project) than the



macro (i.e., program) level. A specific process happened in Uri, which blurred the definition of the program's objectives, and hampered the effectiveness of the program. The tobacco prevention programs are being financed on a matching co-funding basis (canton-Fund). In order to increase the national contribution, Uri sometimes included as cantonal self-contribution activities that bear only little relationship to tobacco prevention. As an effect, this blurred the readability of the program and partly complicated its monitoring and implementation. Whilst the defined output goals—that were little related to smoking prevention—could be achieved, they did not translate into a provable change in the behavior of target groups.

The internal coherence of the program (hypothesis 3) was good but not always optimal. Indeed, cantonal programs are a new form of policy structure, and the cantons need time to adapt to these emerging processes. Our case study highlighted inconsistencies between output and outcome objectives within some projects. The insufficient causality between these two categories of objectives hampered the success of these projects. Secondly, inconsistencies regarding the links between the outcome objectives (i.e., projects) and the impact objectives (i.e., program) were also identified. This was especially the case in regard to the target group smokers. These policy failures constituted a basic incoherence which also impacted the efficiency of the program. In contrast, the outcome–impact chain concerning youth was compact and coherent, thus providing an excellent basis for a successful implementation. However, our case study also showed that a dense range of measures aimed at one target group runs

the risk of duplications.

Regarding the external program coherence (hypothesis 4), our case study highlighted two main negative factors. Firstly, the neglect of the medical setting constituted a possible lack in matter of smoking prevention. The cantonal program could have benefited from an inclusion of this setting allowing to reach many beneficiaries in a facilitating context. Secondly, the saturation of schools was highlighted: they are approached by numerous external demands, and have to make a selection between competing themes (violence, social network, etc.). It is therefore crucial to achieve a good external coordination with the school services in order to obtain access to this setting. We also observed two important positive factors regarding external program coherence: the building of synergies with other addiction services such as for e.g., drug counseling services (coordination of objectives), and the building of synergies with other health promotion programs, for example in order to activate already existing partner-networks (procedural coordination).

Our case study on Uri supports the four hypotheses on the policy concept. Evidence, precision as well as internal and external coherence are crucial factors of a program's conception. Interestingly, the canton of Uri was more successful at the program than the project level as regards the use of evidence, and the reverse was true as regards the precision of objectives.

### *Organization*

Looking at the canton's implementation structure, the case of Uri confirmed the importance of organizational fit. The main findings can be divided into two groups. Firstly, in accordance with



the predefined hypotheses 5 and 6, the structure on the cantonal level proved to be a determining factor for the effectiveness of the tobacco prevention activities, regarding both the appropriateness and the sustainability of the implementation structure. Secondly, the evaluation results showed that there is a need for an extension of the evaluation model which stresses the importance of the collaboration between cantonal and national actors. Although this level of collaboration was not considered in the initial evaluation model, which only considered the cantonal level, its importance became increasingly evident, which is why the analysis was enriched accordingly.

The first group of findings regards the structure of the program at the cantonal level. We identified beneficial and detrimental factors for the program's effectiveness. Firstly, as beneficial factor, the launch of a cantonal program demonstrated the political interest for tobacco prevention, leading to increased legitimation of the related activities. Secondly, the fact that with the program, the tobacco prevention activities were newly coordinated at the cantonal level enabled the actors to improve their networks and thus to identify and use synergies. Thirdly, the collaboration within the canton was additionally facilitated by the relatively small size of the cantonal administration. Due to the pre-existing closeness of the various administration units, the exchange between the different departments was informal and allowed for simple access to different settings. Fourthly, the program manager's capacity to include central actors such as the police (for the enforcement of tobacco control laws) had immediate effects on the program's effectiveness. In cases of strong collaboration, the activities' performance

was increased due to better access to settings and increased legitimacy. Where collaboration with such central partners remained weak, the activities could often not be implemented as planned and lacked public acceptability. Lastly, the relatively inflexible and restrained budget of the canton in conjunction with the partially unpredictable financing of the Fund led to some financial uncertainties and sustainability problems. Where national funding ceased unexpectedly, the canton could not always be sufficiently flexible and raise the necessary funds to continue the concerned activities, leading to stop-and-go problems and losses regarding effectiveness.

The last point leads us to the second group of findings with regard to the program's structure, namely the collaboration of cantonal and national actors, where additional factors relevant to effectiveness were identified. Along with the above-mentioned uncertainties concerning the national funding of certain prevention activities, other changes in regulations as e.g., reporting procedures led to disruptions in the implementation. Besides the immediate practical implications such as the underfunding of certain activities, the perceived lack of continuity in the procedures caused growing discontent at the cantonal level. In the case of Uri, these problems of cooperation already started while the program was designed and continued in the subsequent approval stage, processes that were lengthy and unpredictable in the view of the canton. In addition to the relatively long preparation phase and the uncertainty concerning the Fund's approval of the cantonal program, these uncertainties led to decreasing political support and growing demotivation of the implementing partners. This rather

unfortunate start of the collaboration had implications on the implementation phase and the dissatisfaction of the main cantonal actors could not be reduced until the program's completion. From the national side, these changes and procedures were aimed at improving the quality of prevention policies in Switzerland. We observe here a conflict between the need for continuity on the one hand, and the wish to redirect prevention policies on the other hand.

In conclusion, the case of Uri demonstrates that whilst studying the program structure, it is not sufficient to consider exclusively the cantonal level. The importance of the organizational structure cannot be assessed without considering the collaboration between the cantonal and the national level. Even if the cantonal structure supports the successful implementation of the program, the overall success of the prevention activities depends on a functioning collaboration between the Fund and the cantonal authorities. In regard to the hypothesis 5, it can be concluded that the appropriateness of the program's structure is decisive for its success. As postulated in the hypothesis 6, the sustainability of the program structure, especially with regard to financial security and monitoring processes, proved to be an additional important factor influencing the program's overall performance.

### *Output*

Hypothesis 7 focuses on the compliance of the program implementation at the output level. The evaluation showed that a vast majority of output objectives were achieved by the implementing partners. However, a closer look at the implemented activities exposed some short-

comings in regard to the services' effects on the target groups. The comparison of the level of achievement of both output and outcome goals indicates that there was in some cases a lack of causality between these two levels of analysis. In fact, a much smaller proportion of outcome targets than output goals were achieved. One of the underlying reasons was insufficiently ambitious output goals, which for example did not go beyond the development of organizational structures which were necessary to provide the actual services. Looking at the causes for non-achievement of output goals, the analysis identified certain management problems, such as an absence or delays of responses to arising issues. For instance, the importance of multipliers was recognized in the program's conception, while their advantages were not fully exploited in the implementation due to unachieved inclusion of these key actors.

In summary, the assessment of the output level requires in a first step a qualitative assessment of the target definitions in order to provide well-informed statements about the level's effectiveness and efficiency. Moreover, in cases of well-defined goals, managerial capacity is decisive for the success or failure of service provision. In general, as the related hypothesis 7 states, the appropriateness of the outputs increases the activities' performance concerning the outcome dimension when the causality between these two levels is given.

### *Outcome and Impact*

We turn now to hypothesis 8 about the effects of the outcome effectiveness as well as to hypothesis 9 about the effects of a "multi-channel" logic on the program effectiveness.

Firstly, as discussed above, a smaller proportion of outcome goals were achieved in comparison with the outputs. Some of the underlying reasons were due to policy failures (such as the incoherency between the two levels as well as the partially problematic definition of output goals). Implementation problems such as recruitment difficulties also led some projects to fail because of a lack of participants. Finally, the small outcome of certain projects that had only little to do with tobacco prevention also hampered the impact logic of the program.

Secondly, the case of Uri provided strong evidence for the importance of the “multi-channel” logic. Target groups which were approached with a well-balanced set of projects in different settings, were more likely to be addressed in an effective manner than groups with a weakly conceptualized approach. More specifically, the first main target group, youth, benefited from an advantageously planned set of projects in the settings school (both compulsory education and post-obligatory education), leisure and sports, as well as market. The subsequent successful implementation of these activities did not only target the adolescence in different surroundings, but also with a balanced selection of behavioral and structural prevention as well as information activities. These projects were designed for different ages within the superordinate target group. By contrast, the second main target group smokers could only be addressed in a more limited way. Already in the program design, the chosen approach included only a very restricted number of projects for smokers and only a small proportion of the available funds were allocated to this group. Besides this not very advantageous basis, the respective projects faced

implementation problems and reached only a very small proportion of their audience. Taken together, smokers only benefited from limited smoking cessation aids offered by the program’s activities due to planning and implementation problems. Overall, the “multi-channel” logic turned out to be a promising approach on the one hand, but also resource intensive on the other hand.

In summary, the case of Uri supports the two initial outcome-related hypotheses 8 and 9. The findings show that the better the outcome performance of the individual projects, the stronger the overall preventive effectiveness of the program. Additionally, activities that follow a well-designed and rigorously implemented “multi-channel” logic (in other words, a coherent variety of different prevention approaches) lead to better effects on the target groups than isolated projects.

### *Context*

Finally, our model also aims at investigating the role of the cantonal context along the three following dimensions: the initial extent of the public health problem (hypothesis 10), the socio-structural context (hypothesis 11), and the political context (hypothesis 12). These three dimensions turned out to be crucial factors for the effectiveness of the program, although not always showing the expected effect.

We first hypothesized that a comparatively large tobacco-related problem in the canton would increase the program’s effectiveness, given the relative higher potential for improvement. In fact, this was not the case in Uri, where a particularly important problem of snuff tobacco negatively impacted the effectiveness of the program. Very few improvements could

be achieved in this regard, and moreover, the activities directed toward snuff rather lowered the overall acceptance for the program within the population. Snuff is a traditionally established practice in the canton of Uri and is associated with local particularism. In this context, prevention regarding snuff was perceived as an external intrusion drove from the capital city without any regard to the local culture. As a result, a larger dissatisfaction with the whole tobacco prevention program arose in the canton, negatively impacting its effectiveness.

Secondly, the socio-structural context negatively affected the program in two respects. On the one hand, the size of the canton and the population were too small for certain activities to be implemented successfully. For example, the canton lacked the critical mass of interested persons to conduct smoking cessation group courses. For this reason, many of the planned activities had to be cancelled. On the other hand, the small size of the canton also rapidly led to a strong feeling of saturation in regard to tobacco prevention. The program was large and the program managers reported certain weariness among the population, who felt overwhelmed by smoking prevention. Finally, as already mentioned, the strong inter-knowledge networks and the short distances between prevention actors allowed for rapid action and were identified as positive socio-structural factors.

Thirdly, our hypothesis regarding the political context was not confirmed, as in the case of Uri the administration managed to launch and implement an important tobacco prevention program in a little experimented canton with minimal political support. In a first step, the prevention entrepreneurs managed to establish some

preventive regulatory standards (passing of a cantonal law), which secured long-term activities in this field. In a second step, the program leaders were able to compensate for the minimal financial and political support at the local level through the national support, obtained as a result of the new cantonal smoking prevention programs.

To sum up, two of our context hypotheses (10 and 12) were not supported by the case study: the initially widespread snuff problem could not comparatively increase the program's impact effectiveness, and the program was successfully implemented with only little political support because the program leaders successfully found complementary support elsewhere. The hypothesis 11 related to the socio-structural context led to a contrasted result, as the small size of the canton had both negative and positive effects.

## Conclusion

In this paper, we presented the evaluation design for a complex multilevel program embracing the federal level, the cantonal program level, and the project level where target groups are directly addressed. We employ Pawson and Tilley's realist evaluation approach in order to do justice to the varying context factors and mechanisms that impact the cantonal programs leading to varying effectiveness. The application of the model to the canton of Uri shows that the numerous vertical and horizontal relations play a crucial role for the program. For example, we saw that with the help of the program as a new form of policy structure, an extensive tobacco prevention policy can be initiated, in spite of a rather moderately favorable political context at the local

level. Similarly, we observed how the reluctance of certain local actors can undermine the implementation of the program, for instance as regards particular cantonal smoking consumption patterns. As a general learning for the evaluation of complex programs, we first state that the inclusion of all related vertical and horizontal relations should be considered and that no monocausal effects can be singled out in programs where multiple interventions address the same problem.

Secondly, the case of Uri strongly implies to go beyond the borders of the actual program organization and include factors that do not directly interfere with the policy delivery as such. In particular, all levels of action affecting a program have to be taken into account. In the case of Uri, a crucial organizational factor was the relationship between the federal level with the Fund as a main actor and the cantonal program coordinator. Problems in cooperation of these two actors arose in various phases of the program, decreasing the local, political, and social support and hampering the implementation. Hence, this finding implies a stronger inclusion of the federal level and, in particular, the history of the collaboration between the national and cantonal actors for outcome and impact evaluation, even if the evaluated programs are cantonal.

Finally, Uri is just the first of at least 14 cantons implementing a tobacco prevention program along the lines of the Fund. These new cantonal programs are currently being evaluated within the frame of our research project. The new tobacco policies resulting from these cantonal programs provide simulating examples of a complex multi-level governance system, combining national and sub-national steering processes. The comparison

of the findings will allow for more robust results as to our hypotheses than could be gathered within this single-case application.

## References

Bala, Malgorzata M., Lukasz Strzeszynski, Roman Topor-Madry, and Kate Cahill. 2013. "Mass Media Interventions for Smoking Cessation in Adults." *Cochrane Database of Systematic Reviews* 6: CD004704.

Balthasar, Andreas, and Franziska Müller. 2014. "Die Verbreitung evidenzbasierter und gleichstellungssensitiver Informationen in den Entscheidungsprozessen kantonaler Steuer- und Sozialtransferpolitik: eine quantitative Analyse." *Swiss Political Science Review* 20 (1): 70–95.

Befani, Barbara, Simone Ledermann, and Fritz Sager. 2007. "Realistic Evaluation and QCA: Conceptual Parallels and an Empirical Application." *Evaluation: The International Journal of Theory, Research and Practice* 13 (2): 171–192.

Blatter, Joachim, Clara Bombach, and Roman Wiprächtiger. 2015. "Enhancing Gender Equity through Evidence-based Policymaking? Theorizing and Tracing the use of Systematic Knowledge in Family and Tax Policy Reforms." *European Policy Analysis* 1 (1): 3–34.

Egeberg, Morten. 2003. "How Bureaucratic Structures Matter: An Organizational Perspective." In *Handbook of Public Administration*, eds. B. Guy Peters, and Jon Pierre. London: Sage, 116–126.



- Howlett, Michael. 2009. "Policy Analytical Capacity and Evidence-based Policy-making: Lessons from Canada." *Canadian Public Administration* 52 (2): 153–175.
- Howlett, Michael, and Jeremy Rayner. 2007. "Design Principles for Policy mixes: Cohesion and Coherence in 'New Governance Arrangements.'" *Policy and Society* 26 (4): 1–18.
- Ismayr, Wolfgang. 2009. "Die politischen Systeme Westeuropas im Vergleich." In *Die politischen Systeme Westeuropas*, ed. Wolfgang Ismayr. Wiesbaden: VS Verlag für Sozialwissenschaften, 9–64.
- Kissling-Näf, Ingrid, and Sonja Wälti. 2007. "The Implementation of Public Policies." In *Handbook of Swiss Politics*, eds. Ulrich Klöti, Peter Knoepfel, Hanspeter Kriesi, Wolf Linder, Yannis Papadopoulos, and Pascal Sciarini. Zürich: NZZ, 501–524.
- Knoepfel, Peter, Frédéric Varone, Werner Bussmann, and Luzius Mader. 1997. "Evaluationsgegenstände und Evaluationskriterien." In *Einführung in die Politikevaluation*, eds. Werner Bussmann, Ulrich Klöti, and Peter Knoepfel. Basel, Frankfurt a.M.: Helbing & Lichtenhahn, 78–118.
- Kübler, Daniel. 2000. *Politique de la drogue dans les villes suisses entre ordre et santé. Analyse des conflits de mise en œuvre*. Paris: L'Harmattan.
- Kübler, Daniel, and Sonja Wälti. 2001. "Gouvernance métropolitaine et démocratie: l'exemple des politiques de lutte contre la drogue." *Swiss Political Science Review* 7 (2): 1–25.
- Ledermann, Simone, and Fritz Sager. 2009. "Problem erkannt, aber nicht gebannt. Der Nutzen einer Verknüpfung von Konzept- und Umsetzungsevaluation am Beispiel der schweizerischen Strategie 'Migration und Gesundheit.'" *Zeitschrift für Evaluation* 8 (1): 7–25.
- Linder, Wolf. 1987. *Politische Entscheidung und Gesetzesvollzug*. Bern: Haupt.
- Linder, Wolf. 2010. *Swiss Democracy. Possible Solutions to Conflict in Multicultural Societies*. New York: Palgrave.
- Lund, Johan, and Leif Edvard Aarø. 2004. "Accident Prevention. Presentation of a Model Placing Emphasis on Human Structural and Cultural Factors." *Safety Science* 42: 271–342.
- Mavrot, Céline. 2012. "The Status of Ideas in Controversies on Public Policy. Analyzing Beliefs as Dependent Variables: A Case study on Harm Reduction Policies in Switzerland." *German Policy Studies* 8 (1): 113–156.
- Mavrot, Céline, and Fritz Sager. 2014. "Smoking Prevention Policies in Switzerland: Administrative Coalition and Policy Shifts." Presented at the *23rd World Congress of Political Science*, Montreal, Canada, July 19–24.
- May, Peter J., Joshua Sapotichne, and Samuel Workman. 2006. "Policy Coherence and Policy Domains." *Policy Studies Journal* 34 (3): 381–403.
- Moser, Ladina, and Fritz Sager. 2015. "Pfadabhängigkeit in der Mehrebenensteuerung: Das Beispiel Palliative Care." *Swiss Political Science Review* 21 (3): 437–459.



- Nicholson-Crotty, Sean, and Laurence J. O'Toole Jr. 2004. "Public Management and Organizational Performance: The Case of Law Enforcement Agencies." *Journal of Public Administration Research and Theory* 14 (1): 1–18.
- Patton, Michael Q. 1997. *Utilization-Focused Evaluation: The New Century Text*. London, Thousand Oaks, New Delhi: Sage Publications.
- Pawson, Ray. 2002a. "Evidence-based Policy: In Search of a Method." *Evaluation* 8 (2): 157–181.
- Pawson, Ray. 2002b. "Evidence-based Policy: The Promise of 'Realist Synthesis'." *Evaluation* 8 (3): 340–358.
- Pawson, Ray, and Nick Tilley. 1997. *Realistic Evaluation*. London, Thousand Oaks, New Delhi: Sage Publications.
- Rossi, Peter H., Howard E. Freeman, and Mark W. Lipsey. 1999. *Evaluation. A Systematic Approach*, Sixth Edition. London, Thousand Oaks, New Delhi: Sage Publications.
- Rothmayr, Christine, Frédéric Varone, and Éric Montpetit. 2003. "Does Federalism Matter for Biopolitics? Switzerland in Comparative Perspective." *Swiss Political Science Review* 9 (1): 109–136.
- RS 641.316. *Ordonnance du 5 mars 2004 sur le fonds de prévention du tabagisme (OFPT)*. Berne: Chancellerie fédérale.
- Sager, Fritz. 2003. "Kompensationsmöglichkeiten föderaler Vollzugsdefizite." *Swiss Political Science Review* 9 (1): 309–333.
- Sager, Fritz. 2004. "Verwaltung, Politik und Wissenschaft in der kantonalen Alkoholprävention." *Sozial-und Präventivmedizin. International Journal of Public Health* 49 (3): 208–215.
- Sager, Fritz. 2007. "Making Transport Policy Work: Polity, Policy, Politics, and Systematic Review." *Policy & Politics* 35 (2): 269–288.
- Sager, Fritz. 2008. "Securing the Long-Term Bases of the Dual System: A Realistic Evaluation of Apprenticeship Marketing Projects in Switzerland." *Journal of Vocational Education & Training* 60 (3): 327–341.
- Sager, Fritz, and Céline Anderegg. 2012. "Dealing with Complex Causality in Realist Synthesis: The Promise of Qualitative Comparative Analysis (QCA)." *American Journal of Evaluation* 33 (1): 60–78.
- Sager, Fritz, and Céline Mavrot. 2015. "Les spécificités méthodologiques de l'évaluation en Suisse." In *Regards croisés sur l'évaluation en Suisse*, ed. Katia Horber-Papazian. Lausanne: PPUR, 175–195.
- Sager, Fritz, Céline Mavrot, Susanne Hadorn, and Lyn Pleger. 2015. *Evaluation des kantonalen Tabakpräventionsprogramms Uri 2012–2015*. Bern: Kompetenzzentrum für Public Management der Universität Bern, 131p.
- Sager, Fritz, and Christian Rüeßli. 2005. "Die Evaluation öffentlicher Politiken mit föderalistischen Vollzugsarrangements." *Swiss Political Science Review* 11 (2): 101–129.

Sager, Fritz, and Christine Zollinger. 2011. "The Swiss Political System in a Comparative Perspective." In *Switzerland in Europe*, eds. Christine Trampusch, and André Mach. London, New York: Routledge, 27–42.

Sager, Fritz, and Markus Hinterleitner. 2014. "Evaluation." In *Lehrbuch der Politikfeldanalyse*, eds. Klaus Schubert, and Nils C. Bandelow. München, Wien: De Gruyter/Oldenbourg, 437–462.

Sager, Fritz, and Yvan Rielle. 2013. "Sorting Through the Garbage Can. Under What Conditions Do Governments Adopt Policy Programs?" *Policy Sciences* 46 (1): 1–21.

Vatter, Adrian. 2007. "Federalism." In *Handbook of Swiss Politics*, eds. Ulrich Klöti, Peter Knoepfel, Hanspeter Kriesi, Wolf Linder, Yannis Papadopoulos, and Pascal Sciarini. Zürich: NZZ, 77–99.