POLITICS
AND
PRACTICES
OF

INTERGOVERNMENTAL
EVALUATION

OLAF RIEPER
AND
JACQUES TOULEMONDE
EDITORS

WITH A FOREWORD BY
RAY C. RIST

**POLITICS
AND
PRACTICES
OF**

**INTERGOVERNMENTAL
EVALUATION**

Comparative Policy Analysis Series

Ray C. Rist, Series Editor

Program Evaluation and the Management of Government,
edited by Ray C. Rist

Budgeting, Auditing, and Evaluation,
edited by Andrew Gray, Bill Jenkins, and Bob Segsworth

Can Governments Learn?
edited by Frans L. Leeuw, Ray C. Rist, and Richard C. Sonnichsen

Politics and Practices of Intergovernmental Evaluation,
edited by Olaf Rieper and Jacques Toulemonde

POLITICS
AND
PRACTICES
OF

INTERGOVERNMENTAL
EVALUATION

OLAF RIEPER
AND
JACQUES TOULEMONDE
EDITORS

WITH A FOREWORD BY
RAY C. RIST

First published 1997 by Transaction Publishers

Published 2017 by Routledge

2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

711 Third Avenue, New York, NY 10017, USA

Routledge is an imprint of the Taylor & Francis Group, an informa business

Copyright © 1997 by Taylor & Francis

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

Notice:

Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Library of Congress Catalog Number: 96-30780

Library of Congress Cataloging-in-Publication Data

Politics and practices of intergovernmental evaluation / edited by Olaf Rieper and Jacques Toulemonde ; with a foreword by Ray C. Rist.

p. cm. — (Comparative policy analysis series)

Includes bibliographical references and index.

ISBN 1-56000-256-5 (alk. paper)

1. Central-local government relations—Evaluation. 2. Federal government—Evaluation. 3. Political planning—Evaluation. I. Rieper, Olaf. II. Toulemonde, Jacques. III. Series.

JS113.P64 1996

351.09'3—dc20

96-30780

CIP

ISBN 13: 978-1-56000-256-7 (hbk)

Contents

Foreword	vii
<i>Ray C. Rist</i>	
Introduction: Politics and Practice of Intergovernmental Evaluation	1
<i>Jacques Toulemonde and Olaf Rieper</i>	
1. Evaluation and Autonomy in Sweden	25
<i>Jan Eric Furubo</i>	
2. Evaluative Institutions in England and Wales: Weak Versions of Intergovernmental Evaluation	41
<i>Mary Henkel</i>	
3. Evaluation in the United States: Cooperative But Not Intergovernmental	57
<i>Linda G. Morra</i>	
4. Two Cases in Intergovernmental Evaluation in Canada: "Parallel Play" and Cooperation Without Policy Consequences	75
<i>Robert V. Segsworth and Dale H. Poel</i>	
5. Intergovernmental Evaluation of an EU-Funded Regional Development Program in Denmark	99
<i>Olaf Rieper</i>	
6. Europe and the Member States: Cooperating and Competing on Evaluation Grounds	117
<i>Jacques Toulemonde</i>	
7. Evaluation of the Federal Office of Environmental Protection: Across Two Levels of Government	133
<i>Willi Zimmermann and Peter Knoepfel</i>	
8. Intergovernmental Evaluation: Patterns and Prospects	151
<i>Olaf Rieper and Jacques Toulemonde</i>	
Contributors	167
Index	171

Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Foreword

One of the challenging and intriguing difficulties of research is to recognize that which is right in front of you. The research impulse is to seek out the nuances, the idiosyncratic, and the obscure. The self-evident is ignored.

This difficulty has not escaped the evaluation community. Great intellectual effort has gone into the development of sophisticated designs and methodologies for the study of individual policies, programs, and projects. Costly efforts to find the smallest evidence of a policy or program impact are undertaken in the presumption that such data are central to policy decision making. The emphasis has been overwhelmingly to find discrete impacts from discrete governmental initiatives. Indeed, in the United States, the most frequently used term when discussing evaluation is to refer to “program” evaluation, that is, the study of an individual initiative or activity.

What has been ignored in the evaluation community is the fact of the intergovernmental nature of political and policy governance—regardless of which specific nation state one might choose. Whether it is Canada, the United States, England, Denmark, Sweden, Germany, Japan, or any other of the industrial countries, the governmental structure is essentially a web of interrelated policies, programs, and projects. The political mosaic has both vertical and horizontal components, but it is, indeed, a mosaic. Understanding local responsibilities and requirements for funding housing construction, for example, necessitates also an understanding of the role that regional and national governmental agencies and administrations also play in the scheme of housing. The same can be said for transportation, energy, health care, telecommunications, and on and on.

The point, quite simply, is that governance in complex democratic societies necessitates the coordination and cooperation of multiple levels of government. Sometimes, there is duplication and overlapping levels of jurisdiction, at other times it is more neat. But the totality of governing agencies at whatever level are woven together into a political system. It is the nature of this system that it is inherently intergovernmental. There are

areas where one level of government has most if not sole responsibility for a governance function, but these are noteworthy by their being the exception. In their totality, democratic systems embrace local, regional, and national levels of governance; they embrace multiple sectors of public activity and responsibility; and authority to act is portioned out via an intricate set of regulations, laws, judicial rulings, and constitutional requirements.

If the above is at all accurate as a quick sketch of the system of governance in democratic political systems, then the appearance of this volume, *The Policy and Practice of Intergovernmental Evaluation*, merits special attention. It does so for two reasons, each of which would by itself be sufficient justification for the publication. Taken together, this book becomes both the benchmark and landmark for work in the area of the evaluation of intergovernmental policies, programs, and projects.

First, it has taken on directly the issue of intergovernmental evaluation, a topic that seems to have illuded the attention of the evaluation community for the past thirty years. Major textbooks in the field do not address the topic. It is not, for example, listed in the index of either Rossi and Freeman's opus, *Evaluation* (5th edition) or in the index to Patton's volume, *Qualitative Evaluation and Research Methods* (2nd edition). Further, and this is most telling, it is not listed as one of the topics or concepts discussed by Michael Scriven in his recently published, *Evaluation Thesaurus*, a more than 400-page treatment of the key terms in evaluation.

The irony of this should not be lost. Evaluation has grown and matured as a response to the needs of the public sector in general and the policymaking community in particular. What appears to have emerged over these past decades is a situation where the policy community did not ask and the evaluation community did not offer. Both may intrinsically have well understood the difficulty in moving into intergovernmental issues. The policy community has had reason to shy away for in this realm, accountability is diffuse and fractured, the demands for coordination and cooperation are large, and the evidence of successful outcomes and impacts are hard to demonstrate—as the chapters in this present book make clear. Further, it is hard for any individual politician or policy maker to take much credit for what outcomes are achieved. The diffusion of responsibility across levels of government and across multiple actors in the policy arena means any claim to fame is hard to make. From their vantage point in making the political calculations, there are too many players and too little visibility.

The evaluation community has also not stepped forward to take the lead in this area. Intergovernmental programs make for a most difficult set of circumstances for evaluators. To wit: there are often diffuse sets of goals at different levels of government (quite deliberately done to ensure the acceptance of the initiative at the different levels—allowing each level to interpret the intentions to their own advantage); there are different sets of players who have different sets of expectations for the policy or program and will define success in their own terms; the data systems are not aligned across levels of government to ensure some degree of consistency in measurement and tracking of performance; and the range of stakeholders can vary so considerably at different governmental levels that achieving consensus on the goals and objectives of the evaluation itself will inevitably be difficult and maybe impossible. In the face of these circumstances, the evaluators quietly slip back into the comfortable environs of studying individual policies, programs, or projects, and turn away from the complexity of the intergovernmental system.

The second major contribution of this volume lies in it being comparative and cross-national in its perspective and content. No other such book exists on this topic. The material presented here not only provides a systematic theoretical and empirical treatment of intergovernmental evaluation, but does so with case material from seven nations and the European Union. The materials range from the study of individual cantons and their relation to the Swiss government across a range of topics to the manner in which the European Union is now driving more and more of the regional and national policy agendas of the member states. Indeed, the emergence of the European Union as a powerful political force in Europe is directly affecting the development of evaluation as a craft. Questions are now being asked and levels of expertise are being expected across Europe that did not exist even a scant five years ago.

The strengths of this comparative approach are several. The comparative perspective reinforces the intellectual vitality and robustness of the concept of intergovernmental evaluation. The fact that it is not much discussed does not mean it is not present nor important. Many important issues in the policy arena go largely undiscussed. Further, it is a concept that clearly holds across different nation states with different political and policy configurations.

The comparative perspective also brings into sharp relief the manner in which governments have chosen to organize themselves and along what

dimensions in different countries one finds similarities and contrasts. Understanding how, for example, Switzerland and Sweden each organize their intergovernmental systems brings clarity to an understanding of each in comparison to the other. The same would hold true in the instances of Canada and the United States. Indeed, the comparisons need not be in matched pairs. The framework in this book allows one to look across a number of countries on any given dimension and see how each country responds to that particular aspect of governance and how they seek to know of the performance of that dimension via evaluation.

Further, it is in the context of deciding how to evaluate such intergovernmental programs within the different countries that the evaluator has to know the organizational logic of the government. Effective intergovernmental evaluation cannot occur when the evaluator is working with blinders firmly in place, focusing on just the individual policy, program, or project. As noted above, they are webbed into the policy structure of the country. It is this understanding of the web that becomes so crucial to grasping the nature of the intergovernmental relationships.

The editors and contributors to this volume are to be commended for their efforts to bring into sharp relief the central role that intergovernmental relations play in the governance of democratic societies. Further, they have taken us into the uncharted waters of how we might evaluate such relations. That they have done so with such clarity and conviction that intergovernmental evaluation is central to the governance of our societies deserves both our attention and our thanks.

Ray C. Rist
Washington, D.C.

Introduction: Politics and Practice of Intergovernmental Evaluation

Jacques Toulemonde and Olaf Rieper

Intergovernmental Evaluation is an Unknown Land

In almost all countries and policy domains, governments interact in designing, financing, or implementing their programs. This book attempts to demonstrate how these same governments also interact when the time has come to evaluate these programs. The focus here is on how evaluations may be decided, designed, managed, and utilized within an intergovernmental system.

The presumption here is that intergovernmental evaluation (IGE) is of critical interest, and can be meaningfully distinguished from other kinds of evaluation. The reasons why serious attention must be paid to IGE lie in the extensive development of joint policies and programs by multiple levels of government, a trend which is much more pervasive than the reading of constitutions and laws suggest. Consider the following examples.

A look at the federal system in the United States reveals several thousand intergovernmental programs. Lobbyists employed by business firms, labor unions, and citizen groups and the employees of national, state, and local governments constantly cross each other's paths to influence decisions at all levels. The U.S. federal system is ideally seen as a layer cake, but the marble cake is now the standard analogy (Sharkanski, 1981).

In Germany, another federal system, the exercise of most governing functions is shared between the federal government and the *Länder*. These common tasks are of growing importance despite the fact that the constitution establishes rigorous separations between the different levels (Scharpf, 1988).

2 Politics and Practices of Intergovernmental Evaluation

In the French unitary system, the hierarchic model has exploded in the last few decades. This results both from the setting up of the European Union and from the emergence of autonomous regional and local powers. The new levels of government were initially thought of as having independent fields of jurisdiction. In practice, the instances of overlapping are so frequent that intergovernmental policies are becoming the rule. In some extreme but not exceptional cases, European programs involve no less than five levels of government before producing their output on the French territory.

In the smaller, unitary Nordic countries various distinct levels of government have been in existence for more than half a century: the municipalities, the counties and the national governments. Each level has its own elected body of governance and its specific tasks. However, especially during the last decades, national policies have been crossing through the various levels, reflecting a high degree of inter-level dependency.

These examples confirm that many policies and programs involve several levels of government (federations, states, regions, municipalities) and that the sharing of responsibilities between levels is often far from stabilized. Consequently, we should expect to find a number of evaluations which intend to assess joint policies and programs or which aim at justifying new shifts or drifts in the borders of policy fields. But do such evaluations exist? Are they carried out in a manner that reflects the intergovernmental nature of the program? Do they pose specific problems?

Trying to answer these question requires an initial definition of intergovernmental. We propose the following: an evaluation is intergovernmental if two levels of government interact during at least one stage of the evaluation process, the first stage being the decision to evaluate and the final one being the utilization of evaluation findings.

Although our definition is relatively broad, we are surprised to observe that there are very few IGEs in comparison with the worldwide growth of intricate intergovernmental relations. This unbalance seems to be reflected in the available literature. There are many good books and papers on intergovernmental relations, despite Rhodes' (1981) complaint about major remaining defects, but the literature on intergovernmental evaluation is almost nonexistent.

Reasons must be sought for this paradoxical imbalance. IGE obviously faces specific political, organizational, and technical difficulties that will be analyzed in this book and that explain the paradox in part.

However, in a perfect world where all practical difficulties would be overcome, the unbalance might remain, due to the fact that “single-government evaluations” may happen to be the best choice when dealing with joint policies and programs.

Whatever the reasons that make research on IGE so scarce, the authors are convinced that a domain which appears to be unknown territory is worth exploring. However, one cannot progress in new directions with the same means as in secure fields. This explains some of the book’s special features: (1) it starts by taking its bearings firmly and by equipping itself with a set of precise definitions; (2) it gives the authors freedom to explore many different intergovernmental situations, with a deliberate choice in favor of overviewing the whole area in an heuristic way rather than applying a rigorous comparative method; (3) it does not conclude with robust research-based findings, but by marking out new territories.

Before briefly introducing each of the following chapters it should be stressed that the book represents an integrated whole despite the variations in policies, programs, and political contexts. The authors’ ambition is to take the reader on a tour through the landscape of evaluations which has only one predefined aspect, namely their intergovernmental character. In doing so the reader is invited to reflect on each of the national contributions addressing a common set of questions. These questions concern challenges for those who order evaluations as well for those who design and implement evaluations. Some of the questions are framed as hypotheses and expectations, which are dealt with in the following chapters.

One set of challenges is supposed to stem from the very fact that characterizes levels of government, namely that each level of government has its own electorate, its own legitimacy, and therefore potentially has heavily vested interests in the evaluation of a programme, which might differ from those of other levels of government. This introduces a special problem for the evaluator in serving many masters. Obviously a huge and complex pattern of factors influence the surfacing of tension due to accountability to various electorates.

One might expect the political and administrative climate surrounding the program and its evaluation to be of importance. Is the evaluation dealing with very sensitive political questions in an adverse climate, or with routine problems in a consensual climate? Or is it to be expected, that across variations of political climate the type of relations between the levels of government, which are involved in the evaluation process,

4 Politics and Practices of Intergovernmental Evaluation

somehow influence the surfacing of the potential conflict? Does a bilateral relation between government raise more diversity in views than a relation where only one level of government takes the lead? Or in situations where the levels of government do not interact at all? In the next section different types of relations are described.

Thus, in reading the national contributions the reader might ask: Under what circumstances do divergent views develop? Do they materialize in situations where the levels of government have a bilateral and partnership relation to each other? Or also in situations where one level of government is leading the evaluation?

But, however, if conflict is near the surface, how does one cope with it in a more technical sense? By focusing the evaluation to avoid sensitive questions? By narrowing the evaluation questions? By focusing the evaluation on intermediary goals in order to avoid divergent views on outcome in relation to ultimate goals?

The reader is urged to reflect on whether the above challenges of IGE might meaningfully be dealt with in only a technical sense. Does the pressure from different political constituencies of different levels of government make an evaluation relevant to all levels of government almost impossible? And is this the explanation of why so few IGEs are described in the literature in a period where the number of intergovernmental programs explodes?

Another, though related, set of questions concerns the issue of the utilization of IGE. From the challenges mentioned above it is obvious that the tension rooted in the very nature of intergovernmental relations might prevent utilization by more than one level of government or prevent any utilization at all. Following the typology from the literature on the use of evaluation the reader might identify various models of utilization, and ask how the nature of intergovernmental relations hinders or furthers utilization. Does use according to the managerial model become constrained by technical reasons, such as distance from the field data, and difficulties in synthesizing information across levels of government? Is democratic use by politicians and the public restrained by the very nature of IGE, putting the levels of government into competition or even conflict over the evaluation findings? Does a model of shared conceptual use imply a learning feed back loop, become relevant in situations where interlevel consensus is relatively high, and where a policy and professional community exists?

Is it meaningful to distinguish between the different models of use when confronted with the realities of IGE? Is the most promising way to enhance use by securing the closest possible interaction between levels of government?

These are some of the major questions and challenges we believe IGE raises for administrators and politicians ordering evaluations, for those who design and conduct evaluations in an intergovernmental way, and for the researchers dealing with evaluation systems. Trace these issues, and others you may want to raise, during your guided IGE tour through various western countries. The empirical richness of the national chapters will make it worthwhile.

Origins and Content of a Collective Book

This book, and especially this introductory chapter, deals with definitions and tries to identify potentially fruitful paths of research. It is the result of tough and difficult debates that took place over three years within the IIAS Working Group on Policy and Program Evaluation. Chapters 2 to 8 are national contributions.

In our basic set of examples, two cases dealt with the European Community (Rieper and Toulemonde) and one with federal/provincial interaction in Canada. Only the Swedish case was related to a unitary state. For this reason, we made several attempts to add at least one chapter emanating from a centralized state. With this purpose in mind, contacts were made in both France and England and we were very glad that Mary Henkel finally accepted to join our project as she has a long and well-known record of research on topics close to intergovernmental evaluation.

In the light of the striking changes that occurred in the U.S. evaluation landscape in the 1980s, we obviously needed to include an American chapter. In what way was evaluation affected by the decentralization impulse during the Reagan years? Did this lead to more evaluation at state level or to joint evaluation at federal and state level? Linda Morra has accepted to answer these important questions with the authority bestowed on her by the GAO, although she naturally speaks on her own behalf.

Although we finally conclude that distinguishing federal and unitary states does not really make a difference in intergovernmental evaluation, we initially thought that more input should come from federal systems. For this reason we made several requests in the direction of Germany

and Australia. Finally an answer came from Switzerland, not so big a country nor a continent ! But one must realize how much Switzerland, given its exotic constitutional system, is interesting with respect to interaction between levels of government. The Swiss case, together with that of the European Union, provides us with the most decentralized example of policymaking. We are grateful to Willy Zimmermann and Peter Knoepfel for shedding light on a small territory which may help us to a better understanding of large areas of our problem.

The reader has probably realized that the final set of national chapters result more from opportunity than from the rational ruminations of the editors. This could not be avoided in a project which exclusively depended on the goodwill of volunteers. When we consider that it has been possible to study intergovernmental evaluation in such a variety of different situations, we finally think that good fortune must have presided in the process which has led to the following national chapters :

Jan Eric Furubo, from the Swedish National Court of Audit and from the European Evaluation Society, presents a systematic study of intergovernmental relations in the field of education in his country. His chapter focuses on interaction between the national and municipality levels. It presents a very curious and interesting situation where decentralization leads to decreasing interaction, while evaluation seems to concentrate on the remaining intergovernmental relations.

Mary Henkel, from Brunel University, United Kingdom, draws a portrait of evaluation in her country with examples from the fields of social services and education. She focuses on interaction between national and regional levels of government. Her chapter highlights the transmission of evaluation culture from higher to lower levels through a series of processes initially designed for control. Common evaluation culture may constitute as critical a component for IGE as, for example, common goal clarification, common data collection, or common utilization of results.

Linda Morra, from the United States General Accounting Office, summarizes the history of the separate development of evaluation at federal and state levels. Her examples mainly relate to the fields of education and training programs. She illustrates the hypothesis that is the basis of this book and that can be expressed as follows: strong interaction between governments on policy fields and jointly implemented programs do not automatically lead to intergovernmental evaluation.

In their joint chapter *Dale Poel and Robert Segsworth*, from, respectively, the Dalhousie and Laurentian Universities in Canada, present an interesting and rigorous comparison between two IGE cases studies: one in the field of employment and the other in the field of justice. They focus on interaction between federal and provincial levels. Both examples involve strong interaction on policy fields, but the two evaluations take very different forms, one being carried out with a research design that reflects the intergovernmental character of the program and the other unilaterally.

Olaf Rieper, from the Institute of Local Governments Studies in Denmark, has developed an in-depth case study in the field of regional development policy. His chapter involves no less than four levels of government, encompassing Europe, national government, county and municipalities. The program under evaluation results from an intricate sharing of responsibilities between levels. Evaluation itself is formally organized in an intergovernmental fashion. This chapter perhaps shows the situation which is closest to an ideal IGE. As such, it highlights the problems that arise as a result of a multilevel evaluation and suggests ways of solving them.

Jacques Toulemonde, from the ENTPE in France and from Centre for European Evaluation Expertise, presents an overview of evaluation practice within the European Commission, with examples taken in regional policy and R&D policy. He focuses on interaction between European and Member State levels. Various situations are studied, from the strongest to the weakest intergovernmental policies and evaluations. When looking at an environment where the sharing of responsibilities is subject to permanent debates between levels, it appears that some evaluations serve as arguments or even as weapons in the battle for redefining policy borders.

Willy Zimmermann and Peter Knoepfel, respectively from the Universities of Geneva and Lausanne, Switzerland, report on an evaluation they carried out in the field of environment policy. Their work involved confederate and cantonal levels with strong interaction evident for both the policy and its evaluation. This chapter suggests that vertical separations between policy networks, between professional groups or between lobbies may sometimes make more sense than horizontal limits between levels of government.

The final chapter faces the difficult task of synthesizing a deliberately heterogeneous series of contributions. Since the book's exploratory character must be admitted, it cannot claim to draw robust conclusions from a comparison of research materials collected. The most which can be

done is to sketch a sort of geographic map of this rather unknown land whose name is intergovernmental evaluation.

Governments Have Autonomous Political Resources

What are the basic features which distinguish a government from the other type of organizations involved in public activities? Due to the ups and downs of decentralization, the growing trend towards contracting out and the continuing role played by semi-state bodies, it is often difficult to distinguish governments from quasi-governmental organizations, agencies, or mere layers of administration.

Rhodes (1981) suggests that the very nature of a government must be sought in its political autonomy. More precisely, he speaks of discretion in the use of power resources. These power resources are of a various nature: a government may or may not have its own constitutional independence, its own legitimacy, its own political life with specific democratic arenas, its own system of checks and balances, its own public opinion, its own political elites, its own tax-raising power, its own budget, its own administrative skills and expertise, its own information system, its own procedures and controls, and so on. Note that certain power resources are similar in all organizations, governmental or not, for example: budget, control and information. But other resources which belong to the political sphere are specific to governments, for example: legitimacy, public opinion, checks and balances.

Only fully independent democratic nations have governments which enjoy total autonomy in the use of the full range of power resources. If complete autonomy were the critical criterion, it is difficult to accept that provinces, regions, counties or municipalities have governments. But for the purpose of this book, it is more appropriate to speak of relative autonomy and to use Rhodes' criteria for diagnosing how close the institutions under study are to the absolute definition of a government. Obviously, an institution that has no discretion in the use of political power resources is not a government. We believe that this restriction applies to organizations like universities, hospitals and agencies which we do not consider to be subject matter for this book.

Any government is, or should be, accountable to its own electorate. Within a single level of government, these electorates are in juxtaposition. Constituencies are independent of one another and do not overlap

from a geographic point of view. No voter may simultaneously belong to several areas of authority of the same level.

A conceptual problem arises when several governments of the same level decide to cooperate and set up common procedures and institutions for this purpose. In this case, their common group does not constitute a separate level of government as long as its political legitimacy comes from the founding governments. According to our definition, a new level only appears when the cooperation group gains some kind of direct political legitimacy without the intervention of its founding governments.

Consequently, we consider that the powerful associations of local authorities in the Nordic countries belong to the local level, because their democratic basis is local, even if they enjoy political visibility at the national level.

The case of the European Union is less clear since most of the power belongs to the member states although some elements of autonomous political legitimacy do exist at the European level. The European Council of Heads of States and Governments can be viewed as belonging to the national level, especially when it decides under the rule of unanimity. But the same institution may appear to belong to the European level in other circumstances. Once again, the definition turns out to be a question of degree rather than one of clear-cut distinctions.

A common problem arises from the fact that upper levels of government often have to implement their policies at street level. This means allocating money to individual beneficiaries and applying regulations to local citizens, local companies or local groups. This is the policy law of gravity described as follows: "the buck drops down to local officials at street level." The public money and regulations may reach the local level either directly or through intermediary levels of government. If it is via this latter route, one can conceive of a policy fountain with funds cascading down between levels.

At any level of government full responsibility for a policy or program does not automatically mean direct central administration. In fact, there are three ways of implementing a single-government policy: direct central administration, direct decentral administration and indirect administration.

Direct central administration is the simplest scheme. It means that the officials at one level of government are able to implement the policy themselves. They play the role of the "street-level bureaucracy." The European Competition Policy is an example in point: about forty Euro-

pean officials control the merging of enterprises throughout Europe. They are mainly involved in high-level contacts with multinational firms. This example shows that upper levels can efficiently implement a policy on their own, provided the main stakeholders operate at the same territorial level. Nevertheless, direct central administration is generally more suited to the lower levels of government, these being closer to street level.

Direct decentral administration is often used when the government is too distant from its target and equips itself with a street-level layer of administration. The street level layer reports to the central government and the system involves guidelines and controls enabling the central government to maintain full responsibility for the policy. There are plenty of examples of policies which operate through local layers of administration: the FBI in the U.S., the French education system, and the labor laws in Denmark, to name three.

Indirect administration occurs when a government has its policy implemented by controlled tiers: agencies, trusts, non-profit making associations, or private firms. Control is exerted either through contracting out a specific task or through a representative who sits on the board of directors.

An example of indirect administration is that of the network linking the 217 local action groups involved in the E.U. Leader program which aims at promoting rural development through supporting innovative local development. The European network is managed by a Belgian non-profit making association selected after a call for tenders.

Intergovernmental Policies Imply Sharing Responsibility

In the case of intergovernmental relations, the tasks of designing, implementing and/or evaluating are shared between different levels of government. Since each level of government is normally independent of the others, their shared responsibility raises specific problems. For example, when different levels of governments are accountable to different electorates, each level may have legitimate but contradictory goals. Thus, shared responsibility has dramatic consequences on major evaluative questions like the clarification of goals, the assessment of effectiveness, or specification as to who is accountable. Since the evaluation patterns are likely to be different in a context of shared responsibility, a clear border must be drawn between full and shared responsibility. Special attention must be paid to the risk of confusing layers of administration, those with full responsibility, and those with shared responsibility.

Intergovernmental relations may digress from partnership and co-operation to competition and open conflict. Competing policies run by different levels of government are sometimes targeted at very similar groups or have closely related objectives. In this case, evaluations may become involved in intergovernmental competition or may even be used as a weapon in the fight for power between levels.

Wright (1978) considers that policy areas may be shaped in three ways: (1) the separated model where levels are functionally autonomous and develop few relations; (2) the overlapping model where the main form of relation is bargaining; and (3) the inclusive model where relations are organized in a hierarchic way. Of course, intergovernmental questions are mainly drawn from the overlapping model. Worldwide observation shows that this model is frequent. It does not only apply to federal or confederal systems but also to unitary states.

To consider that a single intergovernmental scheme applies to the totality of a given political system would be a clear case of oversimplification. In fact, a constitutional landscape often includes different areas where one or another of the three models applies. For example, different policy fields such as education, health or technology may be either separated, overlapping or inclusive. Different stages of the policy process may also follow different models: policy design and implementation may overlap while evaluation is separated or vice-versa. Finally, changing models can apply to the same policy area at different points in time, a hierarchic policy becoming more and more overlapping.

Intergovernmental Relations Involve Symmetry of Power Resources

Dependence in intergovernmental systems is a two-way street and must be conceived from the very beginning as belonging to a symmetry of powers. Each level has discretionary use of certain power resources and depends on its counterpart for other resources. According to Rhodes (1981), interaction between levels can follow three different models: the agent model, the cooperative model and the intergovernmental model which is the only one applicable here.

In the agent model, upper level policymakers are free to organize their policy (or program). They can easily reshape the methods of designing, implementing and/or evaluating their policy. They can withdraw responsibility from local governments and contract the job out to agencies, trusts,

or private firms. They can also choose direct management. Even if several levels of government are involved, the balance of power is clearly in favor of the upper level which can design and redesign the inter-organizational structure. Local levels implement the policies of the upper levels under a supervision system which may include financial dependence and detailed controls. When public organizations interact in such a dissymmetric form, the lower level does not play the role of a government but rather that of a layer of administration or an agency.

In the cooperative model, different area authorities at the lower level decide to undertake collectively a strictly delimited policy. This takes the shape akin to a joint diplomatic agreement into which area authorities feel free to enter or not. No government is developed at the upper level and the balance of power is definitely in favor of the lower level. This is typically the case of many international agencies in their relation to national states, but the model can also apply to lower levels. For example, when neighboring municipalities set up a co-operative body to implement common public investments, they face diplomatic problems which are basically comparable to those of governments running an international agency.

In the intergovernmental model, each level relies on its counterpart(s) to design, to implement and/or to evaluate a common policy. Levels are equals according to constitutional rules and each has discretionary use of one or more key resources. They are interdependent in a symmetric way. It is difficult to tilt the intergovernmental relation because no level has the power to reshape the powers of the other levels.

Restricting the Study to the Basic Interaction Across Two Levels

At first glance, it appears plausible to categorize intergovernmental relations into the three traditional forms of constitutions: federal, confederal, and unitary. We have already illustrated that the reality is much more complicated.

In any constitutional system, intergovernmental relations are heterogeneous, each policy area having its own distinctive characteristics. The architecture of intergovernmental relations is not only complex, it is also dynamic. Some authors have seen a worldwide drift to stronger governments at the higher levels, which is obviously the trend in the American federalism (Sharkanski, 1981) and in the European Community. How-

ever, the opposite can be observed in Canada, Spain, Belgium, France, Italy, and in Nordic countries. Meanwhile, countries like Great Britain and Germany have experienced successive periods of relaxed and increased pressure by one level upon another. Considering such opposite trends, Rhodes (1981) thinks that federalism and decentralization have become virtually synonymous.

How to Work Out a Comparative Study in a Disparate and Changing World?

We think that no full comparison can be made between such highly complex systems as Western democracies. Thus, to enable some limited comparisons, we have chosen to proceed with a focus on a simple comparable unit: interaction between two neighboring levels of government. The rationale for the synthesis with which this book ends, lies in the hypothesis that a few common rules apply to the basic relation between two levels of government whatever the country, the constitutional system, the policy field and the "altitude" above street level.

We accept that such an hypothesis may be too restrictive in some cases. For example, when three, four, or five levels of government are related to one another, our reasoning might not cover specific intergovernmental relations of which a typical case is the bypass occurring when a level of government attempts to create direct links with a level which is not its closest neighbor. This is an excellent way to gain power in intergovernmental relations and this game is played by the European Commission and the regions in a way which is unwelcome to the member states. Nevertheless, we think it is worth concentrating on the dual intergovernmental relationship, with findings which are generally applicable, instead of describing curious intergovernmental configurations much like a butterfly collector.

Unified Interaction or Diversely Shaped Relations

However welcome any relevant simplification may be, we cannot avoid looking more closely at the role of the lower level. Since this level is inevitably made up of several area authorities, there is a natural trend for intergovernmental relations to take a differentiated shape according to the different governments which operate at the lower level. The follow-

ing paragraphs present three shapes of interaction which lead either to unified or to diverse intergovernmental relations.

In the first model of bilateral interaction, activities are not entirely shaped by the upper level, but a special negotiation takes place between the upper level and each of the area authorities at the lower level. Negotiations with different governments at the lower level produce different packages of activities, different systems of implementation, and even give birth to different goals or different evaluation criteria. These differences do not only result from the process of conducting parallel negotiations, but also from the fact that the different area authorities at the lower level do not have an equal bargaining power or similar social needs.

It is difficult to find pure examples of such bilateral interactions since every policy or program invariably has a common unified core. But two European examples are not very far from this theoretical model: (1) the French *Contrats de Plan Etat-Région* which organize an intergovernmental framework of public investments through separate negotiations between the central state and each of the twenty-six regions, and (2) the European Community Support Frameworks which result from separate negotiations between the European Community and each member state and lead to distinctive regional policies throughout Europe.

In the second model, interaction is shaped in a standard way at the upper level. Area authorities at the lower levels mainly exert their power through their decision in favor or against the interaction. Of course, having decided to interact, the lower level can develop the whole range of manoeuvres which is always attached to autonomous actors but which is not specific to intergovernmental relations. Numerous cases of such relationships can be found in almost all countries. For example, the Job Training Partnership Act in the U.S. was designed, funded, and evaluated in a unified way at the federal level. States that chose to be involved at implementation and evaluation stages were not allowed to significantly re-shape the federal frameworks.

In the third model, interaction takes place through a single negotiation process that includes the upper level and all the area authorities at the lower level (either directly or indirectly through their representatives). A good example of triangular relationship is the German regional policy (*Gemeinschaftsaufgabe*) which was designed, monitored, and evaluated under the control of a joint council, of which half the members were representatives of the *Länder*, the other half of the federal level. An other

example from a unitary state is the role played by the municipalities in Nordic countries vis-à-vis the national government. The national association of local governments in each country negotiates with various sections of the national government on behalf of all the lower level area authorities. These negotiations end up as genuine treaties between levels of government.

Intergovernmental Evaluation

For the purpose of international comparison the studies presented in this book encompass more than the evaluation of programs. Public management through programs is a somewhat marginal practice in certain countries. We speak more generally of policy evaluation which also includes schemes, laws, contracts, measures, as well as programs.

For the same reasons, we speak of ex-post facto and ongoing evaluation or more generally, of any attempt to judge a policy through a specifically designed collection of field data.

In the same spirit of generalization we allow our studies to encompass the whole scope of evaluation questions: effectiveness, efficiency, coherence, relevance, utility and all goals: measurement, understanding, judgment, and negotiation.

Although our view of evaluation encompasses a lot of different practices, it still fits definitions which are commonly accepted like that of Rist (1989): "Evaluation is the application of systematic research methods to the assessment of policy and program design, implementation and effectiveness."

According to Simeon (1972), an intergovernmental policy is made by independent actors at different levels, who belong to a common social and institutional environment, who share some goals, who negotiate on a set of issues in a nonhierarchical way, who use varying power resources in a strategic and tactical way and finally, who produce common outcomes with consequences for themselves, for others and for the whole system. We do not restrict ourselves to the policies which correspond to this definition. We think that a policy can be intergovernmental at the design stage, at the implementation stage, at the evaluation stage, or at one or more of these stages.

This is why it may be helpful to draw an intergovernmental profile of the specific policy under evaluation. This profile should include, at the very least, the following items:

- Design stage
 1. At which level is the legal framework designed? Are the domains of jurisdiction clearly defined and allocated to different levels?
 2. At which level is the political agenda fixed? Are the dates and deadlines of both levels the same?
 3. Are there discrepancies in the beneficiary areas?
 4. Which model does the decision-making process follow (agent, cooperative, intergovernmental)? Is the policy unified or diverse?
- Implementation stage
 5. Are there co-financed activities involving several different levels?
 6. Is there a common monitoring committee or any other system for jointly controlling the implementation stage?
 7. Is there a joint-information system?
- Evaluation stage
 8. Is there interaction between levels at the evaluation stage, or are there competing evaluations by both levels or a single evaluation by one level?

Figure 1.1 shows an example of a policy profile from the European Framework Program for Research and Development, which has been the subject of an evaluation in 1992, which was an intergovernmental evaluation involving the Member States and the European Commission.

Figure 1.2 gives a second example of an intergovernmental profile of an evaluation: the evaluation of the European "support framework" in Ireland which was undertaken jointly in 1989 by governments at two levels.

Figure 1.1
Intergovernmental Profile of the European Framework
Program for Research and Development

	Member States	Both	European Commission
Design			
Legal framework		x	
Agenda		x	
Decision		x	
Implementation			
Funding			x
Monitoring		x	
Evaluation		x	

Figure 1.2
Intergovernmental Profile of the 1989 Evaluation of the
European "Support Framework" in the Republic of Ireland

	Republic of Ireland	Both	European Commission
Which level?			
Sets up the legal framework		x	
Sets up the agenda			x
Launches the evaluation		x	
Pays for the evaluation			x
Words the questions		x	
Designs the method		x	
Collects and processes data	x		
Receives the draft report			x
Receives the final report		x	
Uses the findings		x	

Is Intergovernmental Evaluation Workable?

In an intergovernmental situation, several technical problems are likely to occur during an evaluation, such as the following.

1. Specific Problems in Collecting and Pooling Data

To a large extent, IGEs encounter the same difficulties as multisite evaluations. Multisite evaluations deal with programs that are implemented in a similar way over multiple sites. Literature on multisite evaluation tries to sort out and solve problems which arise when synthetic conclusions are to be based on decentralized research designs (Lobosco and Kaufman, 1989; Sinacore and Turpin, 1991).

Multiplying the sites of data collection may be a serious impediment to evaluation. The greater the number of lower level area authorities involved in the evaluation process, the more numerous are the gatekeepers who have to be convinced to give access to information and the more varied are the cultural barriers which have to be overcome. Alternatively, it may be worth involving numerous stakeholders at a lower level in the

evaluation process as this may help to enhance utilization, strengthen validity, general scope, and credibility.

Evaluation staff may be hired and supervised either at the upper or lower level. In the latter case, special attention must be paid to training and coordinating local evaluators in order to set up a common framework for data collection (at least for the common core of data), and check the quality of data. Local evaluators may advocate local programs, especially if naturalistic methods are used since these encourage evaluators to immerse themselves in the local environment.

Data can be processed either at the upper or lower level. In the first case, field data are aggregated at the upper level (pooling by lumping). In the second case, the upper level works with data which have been already averaged at the lower level (pooling by averaging). Both methods give rise to specific problems which emerge from the evaluation design (inductive/deductive, quantitative/qualitative).

2. Specific Problems in Identifying and Sharing Outcomes

Outputs are produced at the point where the policy interacts with its target, whereas outcomes are produced at more distant stages of social processes. Outputs remain more or less in the policymakers' hands while outcomes clearly are not. This distinguishes the fields of management which deals with outputs as opposed to evaluation which deals with outcomes.

In the case of IGE, there is a specific difficulty in distinguishing outputs from outcomes, which may lead to confusing, misleading problems of method. For the upper level, outputs can be produced at the street level, where the policy reaches its target. But street-level interaction is beyond upper-level control and what happens at this stage should, in fact, be called outcomes. Viewed from the top, outputs should be the change in lower-level behavior due to upper-level policy. Thus, IGEs must systematically ask two evaluative questions: (1) Is the policy successful in changing the behavior of one or both levels of government? (2) Is the policy successful in changing the behavior of the target group?

The more additional levels of governments are involved in a policy, the more likely it is that evaluation will devote attention to the impacts on intermediary public bodies. This might lead to an evaluation practice

focusing on bureaucratic and diplomatic processes at the expense of social outcomes. Is it possible to observe such a trend?

When two levels of government implement a shared policy, especially when they co-finance the policy in question, it is likely that evaluators will be required to cut the outcome cake into two parts, each part being a benefit resulting from one government's policy. They will be asked to address questions like: what is the value added by each level of government in terms of effectiveness of the policy, or alternatively, is the cost to each level appropriate to its benefits? It is widely recognized that few evaluators are successful in disentangling the complex network of causes and effects. Since they have to handle this additional design problem, intergovernmental evaluators face a hurdle generally more difficult than is found in single-level evaluation.

3. *Specific Problems in Taking Contradictory Criteria into Account*

A naive question must be asked: does the same evaluation method lead to the same conclusion if applied to different levels of government? The answer is clearly negative, as indicated by the following example. In the field of economic and social development policies, many British evaluations use cost-effectiveness logic and attempt to compute "cost per net job" ratios. Many evaluators agree that such evaluations must take into account the displacement effect: an additional job should not be attributed to a policy outcome if it results from the loss of an equivalent job elsewhere. But how is the term "elsewhere" defined—the rest of the beneficiary area, the rest of the UK or the rest of the EC? The result will obviously depend on the evaluator's standpoint and the political considerations of the stakeholders.

This means that the same method, applied on behalf of different levels of government, will produce different results. Similar examples of intergovernmental concerns can be observed in many evaluation techniques. In fact, they clarify a problem that is basically not a technical one. There is necessarily a rationale for a level of government to become involved in a policy field. Thus, it would not make sense if two interacting levels of government shared exactly the same rationale. If this were the case, one of the levels would not be able to justify its *raison d'être*. It is unlikely, therefore, that an IGE could produce common technical answers to evaluative questions posed from two different standpoints.

Is Intergovernmental Evaluation Relevant?

As we noted at the beginning of this chapter, the literature on IGE seems to be remarkably rare. One of the few attempts to discuss the subject (Rosenthal, 1984) comments on the managerial aspects of IGE but concludes that "the most realistic and valuable metaphor for evaluation of intergovernmental programs is that this activity itself become part of a broader societal learning system." We could therefore use two models (managerial and societal learning) to explain the utility of IGE and finally to assess the relevance of this sort of evaluation. In other words, these two models could be called social engineering and social enlightenment. However, we need to add a third model of utilization which applies to evaluation in general and which poses special problems in the case of IGE: this third mode can be termed democratic, (Duran and Monnier, 1992). In this third model, evaluation mainly aims at strengthening the accountability of governments. We will briefly deal with managerial, societal learning, and democratic models in turn.

The first model (managerial) applies to IGE only when the policy under evaluation is itself intergovernmental. The central evaluation question is worded as follows: to what extent is the intergovernmental policy or program a successful one? This model focuses on information feedback aimed at controlling implementation and marginally improving the design.

Managerial use of evaluation in intergovernmental environments faces serious problems since misunderstanding and misinterpreting the meaning of data will likely increase the further the analyst moves away from the source of information, and the more disparate the data sources become. Intergovernmental relations are paradoxical: they call for extensive information feed back but pragmatic constraints means that the data are of little value (Rosenthal, 1984).

The diversity of partners is likely to result in poor comparability. Time schedules and beneficiary areas may not be coherent, statistical definitions may have different meanings, and last but not least, different values may produce shifts in the goals assessed, the categories of impacts or the judgment criteria.

A further difficulty of using IGEs at the higher level becomes apparent when the time comes for synthesizing different local evaluations. In the case of European Structural Policies, the task is so difficult that the upper level has begun to launch more and more single-government evaluations

and relies on them rather than on synthesizing conclusions drawn from IGEs. When confronting the problems of information feedback, we can ask whether intergovernmental evaluation is in fact an appropriate choice.

In the second model (societal learning), the intergovernmental evaluation becomes part of a broad societal learning system rather than a one-shot, clear-cut assessment of success or failure (Rosenthal, 1984). This model applies, to some extent, when a community with a common policy exists. The basic assumptions underlying the policy are presumably shared by the whole society, including politicians and officials at both levels of government. Evaluation is assumed to improve the understanding of the common political system and to transform that system in an incremental way. This function of evaluation is reflected in Karl Popper's perspective of incremental improvements in societal institutions and programs (Popper, 1963). It goes beyond organizational learning, and even beyond governmental learning, which Rist (1993) describes as a part of a web of interacting forces, power systems and institutional arrangements.

It could be claimed that such a societal use of evaluation is likely to develop in contexts where intensity of conflict is modest. This type of evaluation can only lead to utilization when it handles questions which are viewed as debatable by both levels of government.

This model could also be called one of "shared conceptual use." It seems to follow the rule of symmetry that applies to intergovernmental relations, but dissymmetries are observable in the framework. For example, during the Thatcher years, the British government used evaluation openly, especially performance indicators and "value for money audits," as a tool for diffusing neoliberal and managerial ways of thinking in the culture of local governments (Henkel, 1991)

In the democratic model, evaluation is a means to keep society and political personnel informed of policy effectiveness and relevance. In this case, both levels of the government concerned naturally attempt to justify their policies (if policies are competing) or the value they add to the common policy (if policies are intergovernmental). As a consequence, delicate questions can arise such as the extent to which each level contributes to the effectiveness of the policy as a whole or what is the relevance of the outcomes for the electorates at both levels?

In their efforts to provide proof of their respective utility, both levels of government may claim to have produced the same outcomes. If separate evaluations are carried out by each government, they will rapidly

loose their credibility for trying to milk the same cow twice. Consequently, for the sake of credibility there has to be intergovernmental interaction in the evaluation process.

However, launching a common evaluation might raise other difficult problems since such an IGE can quickly show itself to be a zero-sum game. Competing levels are likely to resort to strategies that will, they hope, prove beneficial to themselves: retaining data, undermining the credibility of evaluators, gambling with successive goals, and censoring sensitive questions (Patton, 1986). If accountability is at stake, there must be some intergovernmental features in the evaluation. However, it is likely that interaction between levels will become weaker as the evaluation gets closer to the public arena, for high visibility creates more political vulnerability.

A Final Word

It is clear that among the different kinds of evaluation, IGE appears to be the most vulnerable to high political intrigue. Governments are defined by their autonomy in the use of political resources. Intergovernmental relations are defined by a symmetry in powers. IGE implies certain political agreement on agendas, terms of reference, evaluation methods, intended use of the findings, and so on. But this is not always so. And when this is the case, on what ground can IGE stand?

References

- Duran, P. and E. Monnier (1992). "Le développement de l'évaluation en France, nécessités techniques et exigences politiques." *Revue française de sciences politiques*, 42/2.
- Henkel, M. (1991). "The New 'Evaluative State.'" *Public Administration*, 69/1, 121-36.
- Lobosco, A. and J. Kaufman (1989). "Working with State and Local Service Delivery Systems: The Politics of Evaluating Educational Opportunity Programs at the Community College Level." *Evaluation Review*, 13/2, 141-56.
- Patton, M.Q. (1986). *Utilization-Focused Evaluation*. Newbury Park, CA: Sage.
- Popper, K.R. (1963) *Conjectures and refutations*. London: Routledge and Kegan Paul; New York: Basic Books.
- Rhodes, R.A.W. (1981). *Control and Power in Central-local Government Relations*. Aldershot: Gower Publishing Company. This book includes a comprehensive international bibliography covering the 1960s and the 1970s.
- Rist, R.C. (1989). "Managing of Evaluations or Managing by Evaluations: Choices and Consequences." In R.C. Rist, ed., *Program Evaluation and the Management*

- of Government : Patterns and Prospects across Eight Nations*. New Brunswick, NJ : Transaction Publishers.
- Rist, R.C. (1993) "The Preconditions for Learning: Lessons from the Public Sector." In F.L. Leeuw, R.C. Rist, and R.C. Sonnichsen, eds., *Can Governments Learn? Comparative Perspective on Evaluation and Organizational Learning*. New Brunswick, NJ: Transaction Publishers.
- Rosenthal, S.R. (1984). "New Directions for Evaluating Inter-governmental Programs." *Public Administration Review*, 44/6, 469-76.
- Scharpf, F. (1988). "The Joint-Decision Trap: Lessons from German Federalism and European Integration." *Public Administration*, 66/Autumn, 239-78.
- Sharkanski, I. (1981). "Intergovernmental Relations." In C. Nystrom and W. Starbuck, eds., *Handbook of Organizational Design*. Oxford: Oxford University Press.
- Simeon, R. (1972). *Federal-Provincial Diplomacy: The Making of Recent Policy in Canada*. Toronto: University Press.
- Sinacore, J.M. and R.S. Turpin (1991). "Multisite Evaluations: A Survey of Organizational and Methodological Issues." *New Directions for Program Evaluation*, no. 50.
- Wright, D.S. (1978). *Understanding Inter-governmental Relations*. North Scituate, MA: Duxbury Press.

Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

References

Introduction: Politics and Practice of Intergovernmental Evaluation

Duran, P. and E. Monnier (1992). "Le développement de reevaluation en France, n6cessit6s techniques et exigences politiques." *Revue frangaise de sciences politiques** 42/2.

Henkel, M. (1991). "The New 'Evaluative State.*'" *Public Administration*, 69/1, 121-36.

Lobosco, A. and J. Kaufman (1989). "Working with State and Local Service Delivery Systems: The Politics of Evaluating Educational Opportunity Programs at the Community College Level." *Evaluation Review*, 13/2, 141-56.

Patton, M.Q. (1986). *Utilization-Focused Evaluation*. Newbury Park, CA: Sage.

Popper, K.R. (1963) *Conjectures and refutations*. London: Routledge and Kegan Paul; New York: Basic Books.

Rist, R.C. (1989). "Managing of Evaluations or Managing by Evaluations: Choices and Consequences." In R.C. Rist, ed., *Program Evaluation and the Management o f Government: Patterns and Prospects across Eight Nations*. New Brunswick, NJ : Transaction Publishers.

Rist, R.C. (1993) "The Preconditions for Learning: Lessons from the Public Sector." In F.L. Leeuw, R.C. Rist, and R.C. Sonnichsen, eds., *Can Governments Learn? Comparative Perspective on Evaluation and Organizational Learning*. New Brunswick, NJ: Transaction Publishers.

Rosenthal, S.R. (1984). "New Directions for Evaluating Inter-governmental Programs." *Public Administration Review*, 44/6, 469-76.

Scharpf, F. (1988). "The Joint-Decision Trap: Lessons from German Federalism and European Integration." *Public Administration*, 66/Autumn, 239-78.

Sharkanski, I. (1981). "Intergovernmental Relations." In C. Nystrom and W. Starbuck, eds., *Handbook o f Organizational Design*. Oxford: Oxford University Press.

Simeon, R. (1972). *Federal-Provincial Diplomacy: The Making o f Recent Policy in Canada*. Toronto: University Press.

Sinacore, J.M. and R.S. Turpin (1991). "Multisite Evaluations: A Survey of Organizational and Methodological Issues." *New Directions for Program Evaluation*, no. 50.

Wright, D.S. (1978). *Understanding Inter-governmental Relations*. North Scituate, MA: Duxbury Press.

1 Evaluation and Autonomy in Sweden

Furubo, J.E. (1994). "Learning from Evaluations: The Swedish Experience." In F. Leeuw, R.C. Rist, R. Sonnichsen, eds., *Can Governments Learn? Comparative Perspectives on Evaluation & Organizational Learning*. New Brunswick, NJ: Transaction Publishers.

Larsson, T. (1992). Draft version of forthcoming book "Statssicket-Teori, Praktik och Problem."

Lundh, S. (1994). Memorandum about the evaluation-activities in the National Agency for Education.

Sandahl, R. (1993). "Connected or Separated. In A. Gray, W. Jenkins, and R. Segsworth, eds., *Budgeting, Auditing, Evaluating: Functions and Integration in Seven Governments*. New Brunswick: Transaction Publishers.

Riksrevisionsverket (1993). Memorandum about evaluations in the local authorities.

Riksrevisionsverket (1994). Skolverket—central myndighet i ny roll.

2 Evaluative Institutions in England and Wales: Weak Versions of Intergovernmental Evaluation

Audit Commission (1986). "Making a Reality of Community Care." London: HMSO.

Audit Commission (1992a). "Community Care: Managing the Cascade of Change." London: HMSO.

Audit Commission (1992b). "The Community Revolution: Personal Social Services and Community Care." London: HMSO.

Audit Commission (1992c). "Getting in on the Act: Provision for Pupils with Special Needs: the National Picture." London: HMSO

Audit Commission (1992d). "Getting the Act Together: Provision for Pupils with Special Needs." London: HMSO

Audit Commission (1993). "Report and Accounts." London: HMSO.

Brooke, R. (1989a). The Enabling Authority—Practical Consequences, *Local Government Studies*, 15 (5).

Brooke, R. (1989b). *Managing the Enabling Authority*. London: Longman.

Carter, N. (1991). "Learning to Measure Performance: The Use of Indicators in Organizations." *Public Administration*, vol. 69, no.1 (Spring).

Cave, M., S. Hanney, M. Kogan, and G. Trevett (1988). *The Use of Performance Indicators in Higher Education: A Critical Analysis of Developing Practice*. London: Jessica Kingsley Publishers.

Cm 1599 (1991). *The Citizen's Charter*. London: HMSO.

Cmnd 4814 (1971). *The Organization and Management of Government R & D*. London: HMSO (Rothschild Report).

Cmnd 5406 (1972). *A Framework for Government Research and Development*. London: HMSO.

Cmnd 8616 (1982). *Efficiency and Effectiveness in the Civil Service*. London: HMSO.

Cochrane, A. (1993). *Whatever Happened to Local Government?* Buckingham: Open University Press.

Cordingley, P. and M. Kogan (1992). *In Support of Education*. London: Jessica Kingsley Publishers.

Doig, B. and J. Littlewood (eds.) (1992). "Policy Evaluation: The Role of Social Research." London: HMSO.

Garrett, J. (1972). *The Management of Government*. Pelican.

Gray, A. and M. Jenkins (1982). "Policy Analysis in British Central Government: The Experience of PAR." *Public Administration*, vol. 60, no.4 (Winter).

Henkel, M. (1991a). *Government, Evaluation and Change*. London: Jessica Kingsley.

Henkel, M. (1991b) "The New 'Evaluative State.'" *Public Administration*, vol.69, no.1 (Spring).

Kogan, M. and M. Henkel (1983). *Government and Research: The Rothschild Experiment in a Government Department*. London: Heinemann Educational Books.

Pollitt, C. (1986). "Beyond the Managerial Model: The Case for Broadening Performance Assessment in Government and the Public Service." *Financial Accountability and Management*, 2(3).

Pollitt, C. (1990). "Performance Indicators, Root and Branch." In M. Cave, M. Kogan, and R. Smith, eds. *Output and Performance Measurement in Government: the State of the Art*. London: Jessica Kingsley.

Rhodes, G. (1981). *Inspectorates in British Government*. London: Allen and Unwin.

Ridley, N. (1988). *The Local Right: Enabling not Providing*. London: Center for Policy Studies.

Social Services Inspectorate (1989). *Homes are for Living In: A Model for Evaluating Quality of Care Provided and Quality of Life Experienced in Residential Care for Elderly Persons*. London: HMSO.

Social Services Inspectorate (1992). *Concern for Quality: the First Annual Report of the Chief Inspector, Social Services Inspectorate 1991/1992*. London: HMSO.

Social Services Inspectorate (1993). Raising the Standard: The Second Annual Report o f the Chief Inspector, Social Services Inspectorate 1992/1993. London: HMSO.

Stewart, J. (1992) "Guidelines for Public Service Management: Lessons not to be Learnt from the Private Sector." In Carter et al. eds. Changing Social Work and Social Welfare, Buckingham: Open University Press.

3 Evaluation in the United States: Cooperative But Not Intergovernmental

Advisory Commission on Inter-governmental Relations (1981).
In Brief: State and Local Roles in the Federal System.
Washington: ACIR.

Blalock, A. (ed.) (1993). "National Research."* In
Evaluation Forum, Issue 9. Washington, DC: U.S.
Department of Labor.

Brown, J.R. (1988). "State Evaluation in an Executive
Environment." In C.G. Wye and H.P. Hatry, eds. Timely,
Low-Cost Evaluation in the Public Sector, New Directions
for Program Evaluation, 38, 37-47.

Fisher, M. and C. Place (1976). "Summary of Procedures
Employed." In State Evaluation Reports of Title 1
Migrant Programs. Research Triangle Park, NC: Research
Triangle Institute (Project No 224-1235; Contract
300-76-0095).

Hatry, H.P., R.E. Winnie, and D. Fisk (1973). Practical
Program Evaluation for State and Local Government
Officials. Washington, DC: The Urban Institute Press.

Havens, H.S. (1992). "The Erosion of Federal Program
Evaluation." In C.G. Wye and R.C. Sonnichsen, eds.,
Evaluation in the Federal Government: Changes, Trends, and
Opportunities, New Directions in Program Evaluation, 55,
21-27.

Havens, H.S. (1990). The Evolution of the General
Accounting Office: From Voucher Audits to Program
Evaluations. Washington, DC: U.S. General Accounting Of
fice.

Hyde, A.C. and J.M. Shafritz (1979). Program Evaluation in
the Public Sector. New York: Praeger Special Studies.

Jones, R. (1987). "Keeping an Eye on State Agencies." In
State Legislatures. Denver, CO: National Conference of
State Legislatures.

Klemanski, J.S. (1992). Book review of Evaluating Social
Programs at the State and Local Level: The JTPA Evaluation
Design Project. In M. F. Smith and Carl Wisler, eds.,
Evaluation Practice, 13/3, 207-08.

Levine, R.A. (1983). "Contractor Evaluation for the Federal

Government.” In Robert St. Pierre, ed., *Management and Organization of Program Evaluation, New Directions for Program Evaluation*, 18, 67-68.

Linney, J.A. and A. Wandersman (1991). *Prevention Plus III: Assessing Alcohol and Other Drug Prevention Programs at the School and Community Level: A FourStep Guide to Useful Program Assessment*. Washington, DC: U.S. Department of Health and Human Services, Office for Substance Abuse Prevention, DHHS Publication No. (ADM)91-1817.

Mahone, John A. and Lin Corbin-Howerton (1988). “State Evaluation in an Executive Environment.” In C.G. Wye and H.P. Hatry, eds., *Timely, Low-Cost Evaluation in the Public Sector*, *New Directions for Program Evaluation*, 38, 37-47.

Martinson, K. and D. Friedlander (1994). *GAIN: Basic Education in a Welfare-to-Work Program: Executive Summary*. New York: Manpower Demonstration Research Corporation.

Nathan, Richard P. (1988). *Social Science in Government*. New York: Basic Books.

National Conference of State Legislatures (1992). *The Use of Sunset and Program Evaluation in the State Legislatures: Presentation to the Sunset Review Subcommittee, Indiana Legislative Council*. Denver, CO: National Conference of State Legislatures.

Peck, H.I. and S.E. Triplett (1983). “Management and Organization of an Evaluation Unit in a State Education Agency.” In R.G. St. Pierre, ed., *Management and Organization of Program Evaluation, New Directions for Program Evaluation*, 18, 39-47.

Polivka, L. and L.T. Stryker (1983). “Program Evaluation and the Policy Process in State Government: An Effective Linkage.” In E. Chelminsky, *Program Evaluation: Patterns and Directions*. Washington, DC: American Society for Public Administration.

Provus, Malcolm (1971). *Discrepancy Evaluation: For Educational Program Improvement and Assessment*. Berkeley, CA: McCutchan Publishing Corporation.

Richardson, E.L. (1992). “The Value of Evaluation.” In C.G. Wye and R.C. Sonnichsen, eds., *Evaluation in the Federal Government: Changes, Trends, and Opportunities*, New

Directions for Program Evaluation, 55, 15-20.

Rist, R.C. (1990). "The Organization and Function of Evaluation in the United States: A Federal Overview." In Ray C. Rist, ed., Program Evaluation and Management of Government, New Brunswick, NJ: Transaction Publishers.

Rossi, P.H. and H.E. Freeman (1982). Evaluation: A Systematic Approach. California: Sage Publications.

Trask, R.R. (1991). GAO History 1921-1991. Washington, DC: U.S. General Accounting Office, GAO/OP-3-HR

U.S. General Accounting Office (1992). Dislocated Workers: Comparison of Assistance Programs. Washington, DC: U.S. General Accounting Office, GAO/HRD-92-153BR.

U.S. General Accounting Office (1992). Program Evaluation Issues. Washington, DC: U. S. General Accounting Office, GAO/OGC-93-6TR.

U.S. General Accounting Office (1990). Federal-State-Local Relations: Trends of the Past Decade and Emerging Issues. Washington, DC: U.S. General Accounting Office, GAO/HRD-90-34.

U.S. General Accounting Office (1988). Program Evaluation Issues. Washington, DC: U. S. General Accounting Office, GAO/OGC-89-8TR.

U.S. General Accounting Office (1987). Federal Evaluations: Fewer Units, Reduced Resources, Different Studies from 1980. Washington, DC: U.S. General Accounting Office, GAO/PEMD-87-9.

Wargo, Michael J. (1983). "Management of the Evaluation Function Within the Federal Government." In Robert G. St. Pierre, ed., Management and Organization of Program Evaluation, New Directions for Program Evaluation, 18, 5-25.

Wye, Christopher G. (1992). "The Office of Management and Budget: A Continuing Search for Useful Information." In Christopher G. Wye and Richard C. Sonnichsen, eds., Evaluation in the Federal Government: Changes, Trends, Opportunities. New Directions for Program Evaluation, 55, 65-72.

4 Two Cases in Intergovernmental Evaluation in Canada: "Parallel Play" and Cooperation Without Policy Consequences

Auditor General of Canada (1993). Annual Report. Ottawa: Supply and Services.

Abt Associates of Canada (1991). Comprehensive Review and Evaluation of the Certificate Component of the Ontario Legal Aid Plan. Prepared for the Ontario Ministry of the Attorney General and the Federal Department of Justice.

Comptroller General of Canada (1986). The Effects of the Access to Information Legislation on the Program Evaluation Function. Ottawa: Program Evaluation Branch, Office of the Comptroller General of Canada.

De Koven, Harriet L. (1993). "Federal-Provincial Transfers: Which Way From Here?" Policy Options, vol. 14, no. 10, pp. 45-48.

Dyck, Rand (1993). Canadian Politics: Critical Approaches. Scarborough: Nelson Publishers.

Hartle, D. (1990). "Increasing Government Accountability: A Proposal that the Senate Assume Responsibility for Program Evaluation." Proceedings of the Standing Committee on National Finance. 3 May, Appendix NF-23a.

Johnson, A.W. (1992). "Reflections on Administrative Reform in the Government of Canada, 1962-1991: A Discussion Paper." Prepared for Office of the Auditor General, Ottawa.

Mallory, J. (1965). "The Five Faces of Canadian Federalism." In P.A. Crepeau and C.B. Macpherson, eds., The Future of Canadian Federalism. Toronto: University of Toronto Press.

Onyshko, T. (1991). "Secret paper shows public defender system considered." The Lawyers Weekly (December 13, p. 9).

Poel, D.H., P. Aucoin, and R. Conway (1980). Nova Scotia Legal Aid: A Review and Proposal for Evaluation. Halifax: The Nova Scotia Legal Aid Commission and Federal Department of Justice.

Poel, D.H. (1983). The Nova Scotia Legal Aid Evaluation Report: Entering the "Third Generation." Halifax: The

Nova Scotia Legal Aid Commission and the Federal Department of Justice.

Poel, D.H. (1993). "Cooperation and Conflict in the Evaluation of Legal Aid Services in the Canadian Provinces." In J. Hudson and J. Roberts, eds., *Evaluating Justice: Canadian Policies and Programs*. Toronto: Thompson Educational Publishing, Inc., 1993.

Policy Options (1993). *Fiscal Federalism: Debating Canada's Future* (vol. 14, no. 10). Montreal: Institute for Research and Public Policy.

Segsworth, R.V. (1990). "Policy and Program Evaluation in the Government of Canada." In R.C. Rist, ed., *Program Evaluation and the Management of Government*. New Brunswick, N.J.: Transaction Publishers.

Segsworth, R.V. (1992). "Governmental Activity Review in Ontario." In J. Hudson, J. Mayne, and R. Thomlinson, eds. *Action-Oriented Evaluation*. Toronto: Emerson and Wall.

Segsworth, R.V. (1993). "Public Access to Evaluation in Canada." In R. Conner, J. Hudson, J. Mayne, and M. L. Bemelmans-Videc, eds., *Advancing Public Policy Evaluation*. Amsterdam: North Holland.

Smiley, D. (1987). *The Federal Condition in Canada*. Toronto: McGraw Hill Ryerson Limited.

Task Force on Program Review (1985). *Introduction to the Process of Program Review*. Ottawa: Supply and Services.

Thomas, J.F. (1990). "Opening Statement." *Proceedings of the Senate Standing Committee on National Finance* on 17 May. Ottawa.

Treasury Board (1994). *Treasury Board Manual on Review\ Internal Audit and Evaluation*. Ottawa: Treasury Board Secretariat.

Wheare, K.C. (1946). *Federal Government*. London: Oxford University Press.

5 Intergovernmental Evaluation of an EU-Funded Regional Development Program in Denmark

Beyer, Janice M. and Harrison M. Trice (1982). "The Utilization process: A Conceptual Framework and Synthesis of Empirical Findings." *Administrative Science Quarterly*, 27 (1982): 591-622

Cronberg, Taija et al. (eds.) (1991). *Danish Experiments: The Social Construction of Technology*. Copenhagen: New Science Monograph.

Elder, N., A.H. Thomas, and D. Arter (1983). *The Consensual Democracies? The Government and Policies of the Scandinavian States*. Oxford: Martin Robertson.

Heide, H. Ter (ed.) (1990). *Technological Change and Spatial Policy*, Netherlands Geographical Studies 112. Amsterdam, Utrecht: The Hague.

Knudsen, Tim (1992). "A Portrait of Danish State Culture." In M. Kellstrup, ed., *European Integration and Danish Participation*. Copenhagen: Copenhagen Political Studies Press.

Patton, Michael Quinn (1986). *Utilization-Focused Evaluation*. Beverly Hills, CA: Sage Publication.

Rieper, Olaf, Regis Blais, and Anders Larsen (1988). *Information Technology as a Means for Community Development: Telehouses in Scandinavia and Videotex in France*. [PUBLISHER?]

Rieper, Olaf (1991). "Enterprises and Programs for Regional Technological Development." In Tarje Cronberg et al., eds. *Danish Experiments-Social Constructions of Technology*. Copenhagen: The Danish Social Science Research Council, New Social Science Monographs.

Rist, Ray C. (ed.) (1990). *Program Evaluation and the Management of Government*. New Brunswick: Transaction Publishers.

6 Europe and the Member States: Cooperating and Competing on Evaluation Grounds

Bradley, J. and J. Fitzgerald (1992). "The Role of the Structural Funds: Analysis of Consequences for Ireland in the Context of 1992." Policy Research Series, 13. Dublin: Economic and Social Research Institute.

Greschmann, K. (1992). "Le principe de subsidiarity quelles responsabilités à quel niveau de pouvoir dans une Europe intégrée?" In EIPA (ed.) Subsidiarity, défi au changement. Maastricht: EIPA.

Laredo, Ph. and M. Callon (1990). L'impact des programmes communautaires sur le tissu scientifique et technique français. Paris: Centre de Sociologie de l'Innovation.

Monnier, E. et al. (1991). Programme européen d'analyse méthodologique pour revaluation des politiques structurelles: étude de définition. Lyon: ENTPE-CEOPS.

Quermonne, J.L. (1990). "Existe-t-il un module politique européen?" Revue Française de Sciences Politiques, 40/2, 208.

Rieper, O. (1994). "Evaluation of an EU-Funded Regional Development Program in Denmark," this volume.

Sensi, D. and B. Cracknell (1991). Inquiry into Evaluation Practices in the Commission. Luxembourg: Office des publications de la Commission Européenne.

Sidjanski, D. (1989). "La Communauté Européenne: un gouvernement de comités?" Pouvoirs, 48.

Toulemonde, J. (1995). "Peut-on y valuer la subsidiarity? Elements de réponse inspirés de la pratique européenne," forthcoming.

7 Evaluation of the Federal Office of Environmental Protection: Across Two Levels of Government

Kloeti, U. (1987). *Swiss Democracy—Exception or Model?* Zurich: Forschungsstelle für Politische Wissenschaft, University of Zurich.

Knapp, B. (1986). "Étapes du fédéralisme suisse." In *Manuel Systeme politique de la Suisse*, vol. 3, Berne.

Scharpf, F.W. (1988). "The Joint Decision Trap: Lessons from German Federalism and European Integration." *Public Administration*, vol. 66, Autumn 1988, 239-78.