Continue

Alkin, M., & Christie, T. (Eds.). (2023). Evaluation roots. Theory influencing practice (3rd ed.). Guilford Press. Google Scholar Bamberger, M., & Mabry, L. (2019). RealWorld evaluation. Working under budget, time, data and political constraints (3rd ed.). Sage Publications. Google Scholar Birch, L. M., & Jacob, S. (2019). "Deliverology" and evaluation: A tale of two worlds. Canadian Journal of Program Evaluation, 34(2), 303–328. � Google Scholar Campbell, M., & Ng, R. G. (1988). Program evaluation and the standpoint of women. Canadian Review of Social Policy, 22, 41–50. Google Scholar Dubois, C.-A., Champagne, F., & Bilodeau, H. (2011). Historique de l'évaluation. In A. Brousselle, F. Champagne, A.-P. Contandriopoulos, & Z. Hartz (Eds.), L'évaluation: concepts et méthodes (pp. 28–48). Presses de l'Université de Montréal. Google Scholar Forss, K., Lindkvist, I., & McGillivray, M. (2021). Long term perspectives in evaluation. Increasing relevance and utility. Routledge. � Google Scholar Furubo, J. E., & Vestman, O. K. (2011). Evaluation: For public good or professional power? In P. Eliadis, J. E. Furubo, & S. Jacob (Eds.), Evaluation: Seeking truth or power? Transaction Publishers. � Google Scholar Guba, E., & Lincoln, Y. (1989). Fourth generation evaluation. Sage Publications. Google Scholar Hood, S., Hopson, R. K., & Kirkhart, K. E. (2015). Culturally responsive evaluation. Theory, practice, and future implications. In K. E. Newcomer, H. P. Hatry, & J. S. Wholey (Eds.), Handbook of practical program evaluation (4th ed., pp. 281–317). Wiley. � Google Scholar Jacob, S., & Boisvert, Y. (2010). To be or not to be a profession: Pros, cons and challenges for evaluation. Evaluation. The International Journal of Theory, Research and Practice, 16(4), 349–369. � Google Scholar Lawarée, J., Jacob, S., & Ouimet, M. (2020). A scoping review of knowledge syntheses in the field of evaluation across four decades of practice. Evaluation and Program Planning, 79. . (2021). Applying evaluation criteria thoughtfully. OECD Edition. M. (1991). Evaluation thesaurus (4th ed.). Sage Publications. Google Scholar Stockmann, R., Meyer, W., & Taube, L. (2020). The institutionalisation of evaluation in Europe: A synthesis. In R. Stockmann, W. Meyer, & L. Taube (Eds.), The institutionalisation of evaluation in Europe (pp. 483–522). Palgrave Macmillan. � Google Scholar Thomas, V. G., & Campbell, P. B. (2020). Evaluation in today's world. Respecting diversity, improving quality, and promoting usability. Sage Publications. Google Scholar Thomas, V. G., Madison, A., Rockcliffe, F., DeLaine, K., & Lowe, S. M. (2018). Racism, social programming, and evaluation: Where do we go from here? American Journal of Evaluation, 39(4), 514–526. � Google Scholar Page 2Alkin, M., & Christie, T. (Eds.). (2023). Evaluation roots. Theory influencing practice (3rd ed.). Guilford Press. Google Scholar Bamberger, M., & Mabry, L. (2019). RealWorld evaluation. Working under budget, time, data and political constraints (3rd ed.). Sage Publications. Google Scholar Birch, L. M., & Jacob, S. (2019). "Deliverology" and evaluation: A tale of two worlds. Canadian Journal of Program Evaluation, 34(2), 303–328. � Google Scholar Campbell, M., & Ng, R. G. (1988). Program evaluation and the standpoint of women. Canadian Review of Social Policy, 22, 41–50. Google Scholar Dubois, C.-A., Champagne, F., & Bilodeau, H. (2011). Historique de l'évaluation. In A. Brousselle, F. Champagne, A.-P. Contandriopoulos, & Z. Hartz (Eds.), L'évaluation: concepts et méthodes (pp. 28–48). Presses de l'Université de Montréal. Google Scholar Forss, K., Lindkvist, I., & McGillivray, M. (2021). Long term perspectives in evaluation. Increasing relevance and utility. Routledge. � Google Scholar Furubo, J. E., & Vestman, O. K. (2011). Evaluation: For public good or professional power? In P. Eliadis, J. E. Furubo, & S. Jacob (Eds.), Evaluation: Seeking truth or power? Transaction Publishers. � Google Scholar Guba, E., & Lincoln, Y. (1989). Fourth generation evaluation. Sage Publications. Google Scholar Hood, S., Hopson, R. K., & Kirkhart, K. E. (2015). Culturally responsive evaluation. Theory, practice, and future implications. In K. E. Newcomer, H. P. Hatry, & J. S. Wholey (Eds.), Handbook of practical program evaluation (4th ed., pp. 281–317). Wiley. � Google Scholar Jacob, S., & Boisvert, Y. (2010). To be or not to be a profession: Pros, cons and challenges for evaluation. Evaluation. The International Journal of Theory, Research and Practice, 16(4), 349–369. � Google Scholar Lawarée, J., Jacob, S., & Ouimet, M. (2020). A scoping review of knowledge syntheses in the field of evaluation across four decades of practice. Evaluation and Program Planning, 79. . (2021). Applying evaluation criteria thoughtfully. OECD Edition. M. (1991). Evaluation thesaurus (4th ed.). Sage Publications. Google Scholar Stockmann, R., Meyer, W., & Taube, L. (2020). The institutionalisation of evaluation in Europe: A synthesis. In R. Stockmann, W. Meyer, & L. Taube (Eds.), The institutionalisation of evaluation in Europe (pp. 483–522). Palgrave Macmillan. � Google Scholar Thomas, V. G., & Campbell, P. B. (2020). Evaluation in today's world. Respecting diversity, improving quality, and promoting usability. Sage Publications. Google Scholar Thomas, V. G., Madison, A., Rockcliffe, F., DeLaine, K., & Lowe, S. M. (2018). Racism, social programming, and evaluation: Where do we go from here? American Journal of Evaluation, 39(4), 514–526. � Google Scholar It's also important to talk about the two different uses for evidence within government – accountability and learning. Evidence is useful for both and they are equally important but there are often tensions between the two. Accountability is about proving something to others. When we speak about holding government accountable for delivering, what we're asking the government to do is to produce evidence that it has done what it was supposed to. In the context of impact assessment and evaluation, accountability can be about proving to others that a programme was effective or that it works the way its theory of change expects it to.   Learning, on the other hand, is about improving. Here, the consumers of the evidence are often the same ones who are creating it because they want that information to better their policy. The key audience is internal, not external, to those who are responsible for designing and implementing the policy. Accountability and learning are both important and the same evidence can serve both purposes. For example, an impact evaluation can prove the effectiveness of a programme to outsiders but also help insiders learn ways to improve. There might also be times when there are questions that the insiders are interested in that the outsiders aren't, or vice versa, and sometimes they even directly conflict. If evidence is going to be used for accountability purposes, it can make people less open and less willing to learn from the evidence as it means admitting they made mistakes. Read an example of how impact evaluation works alongside accountability and learning from the proposal in the Center for Global Development report When Will We Ever Learn? Improving Lives Through Impact Evaluation.   In the Evidence in Public Policy short course, we explore how things should workwhen approaching policy evaluation, both in terms of the technical details of evaluation as well as ways to use evidence. Since it's worth knowing what an ideal process would look like, we'll also talk about ways in which reality often makes this difficult or even impossible and what you can do to improve your evidence use in an imperfect world. Throughout the course, you will be encouraged to keep in mind the dual purposes of evaluation as you study examples of evidence. These dual purposes will also be important when designing your very own impact evaluation on the final module – the course project. You will need to think about how the evidence you create is going to be used and what that implies for your evaluation design. You can our to find out more about how it can benefit your career.   Imagine launching a national program to improve rural education in India. After a few years, you want to know whether the program is fulfilling its objectives, making a difference, and worth continuing. This is where policy evaluation comes into play. Policy evaluation is an essential part of the policy cycle that assesses the overall effectiveness of national programs in meeting their objectives. It is a critical tool for policymakers, providing them with the necessary feedback to make informed decisions. But what exactly does policy evaluation entail, and why is it so important? Table of Contents Policy evaluation is the systematic assessment of the design, implementation, and outcomes of public policies. Its primary aim is to determine the effectiveness of a policy in achieving its stated goals. According to Wholey et al., policy evaluation is an analytical tool that measures performance. It involves collecting and analyzing data to assess whether the policy is working as intended, identifying any gaps or shortcomings, and providing recommendations for improvement. Wollmann describes policy evaluation as a phase in the policy cycle that provides feedback to policymakers. This feedback is crucial for making decisions about whether a policy should continue, be amended, or scrapped. In essence, policy evaluation is about accountability, learning, and improvement. The purpose of policy evaluation Policy evaluation serves multiple purposes, all aimed at enhancing the effectiveness and efficiency of public policies. Here are some of the key objectives of policy evaluation: Assessing efficiency and effectiveness One of the primary purposes of policy evaluation is to assess the efficiency and effectiveness of a policy. Efficiency refers to the extent to which a policy uses resources (such as time, money, and manpower) in a cost-effective manner. Effectiveness, on the other hand, measures the extent to which a policy achieves its intended outcomes. By evaluating efficiency and effectiveness, policymakers can determine whether a policy is delivering value for money and meeting its objectives. Policy evaluation provides reliable information about the performance of a policy. This information is crucial for making evidence-based decisions. For example, if a policy aimed at reducing air pollution is found to be ineffective, policymakers can use the evaluation findings to redesign the policy or explore alternative solutions. Reliable information also helps in setting realistic targets and expectations for future policies. Clarifying values and goals Evaluation helps clarify the values and goals underlying a policy. By analyzing the outcomes of a policy, evaluators can identify whether the policy aligns with the values and priorities of stakeholders. This process can lead to a better understanding of the policy's intended and unintended consequences, and help in refining the goals to better match the needs and expectations of the target population. Policy evaluation provides insights that inform future actions. Based on the evaluation findings, policymakers can decide whether to continue, modify, or terminate a policy. For instance, if an evaluation reveals that a job training program is successfully increasing employment rates, it may be expanded or replicated in other regions. Conversely, if a policy is found to be ineffective, it may be discontinued, and resources can be redirected to more promising initiatives. Encouraging accountability Evaluation promotes accountability by holding policymakers and implementers responsible for the outcomes of a policy. It ensures that public funds are being used effectively and that policies are delivering the intended benefits to the public. Accountability is essential for maintaining public trust and confidence in government actions. Methods of policy evaluation There are various methods used in policy evaluation, each with its strengths and limitations. Here are some common methods: Quantitative methods Quantitative methods involve the use of numerical data to measure policy outcomes. These methods often include: Surveys: Collecting data from a large number of respondents to assess their experiences and perceptions related to a policy. Experiments: Conducting controlled experiments to evaluate the impact of a policy intervention. Statistical analysis: Using statistical techniques to analyze data and identify trends, correlations, and causal relationships. Qualitative methods Qualitative methods focus on understanding the context and experiences of stakeholders. These methods often include: Interviews: Conducting in-depth interviews with stakeholders to gather detailed insights into their experiences and perspectives. Focus groups: Facilitating group discussions to explore the views and opinions of different stakeholders. Case studies: Analyzing specific instances of policy implementation to understand the factors contributing to success or failure. Mixed methods Mixed methods combine quantitative and qualitative approaches to provide a comprehensive evaluation of a policy. By integrating numerical data with contextual insights, mixed methods offer a more holistic understanding of policy outcomes. Challenges in policy evaluation Despite its importance, policy evaluation faces several challenges, including: Data limitations Access to reliable and relevant data is crucial for effective evaluation. However, data limitations, such as incomplete or outdated information, can hinder the accuracy and reliability of evaluation findings. Complexity of policies Policies often involve multiple components and stakeholders, making it challenging to isolate the impact of a specific policy intervention. Evaluators must carefully design their studies to account for these complexities. Resource constraints Conducting thorough evaluations requires time, money, and expertise. Resource constraints can limit the scope and depth of evaluations, potentially affecting the quality of the findings. Political and organizational factors Political and organizational factors can influence the evaluation process. For example, policymakers may be reluctant to support evaluations that could reveal negative findings, or there may be resistance to change within implementing organizations. Conclusion Policy evaluation is a vital tool for enhancing the effectiveness and efficiency of public policies. By systematically assessing the design, implementation, and outcomes of policies, evaluation provides valuable feedback to policymakers, helping them make informed decisions about the future of a policy. Despite the challenges, the benefits of policy evaluation far outweigh the difficulties, making it an indispensable part of the policy cycle. What do you think? Have you ever encountered a policy that you believe needs evaluation? What changes would you suggest based on your observations? This publication follows a first book (Évaluation : fondements, controverses, perspectives) published at the end of 2021 by Editions Science et Bien Commun (ESBC) with the support of the Laboratory for interdisciplinary evaluation of public policies (LIEPP), compiling a series of excerpts from fundamental and contemporary texts in evaluation (Delahais et al. 2021). Although part of this book is dedicated to the diversity of paradigmatic approaches, we chose not to go into a detailed presentation of methods on the grounds that this would at least merit a book of its own. This is the purpose of this volume. This publication is part of LIEPP's collective project in two ways: through the articulation between research and evaluation, and through the dialogue between quantitative and qualitative methods. Methods between research and evaluation Most definitions of programme evaluation articulate three dimensions, described by Alkin and Christie as the three branches of the "evaluation theory tree" (Alkin and Christie 2012). These are the mobilisation of research methods (evaluation is based on systematic empirical investigation), the role of values in providing criteria for judging the intervention under study, and the focus on the usefulness of the evaluation. The use of systematic methods of empirical investigation is therefore one of the foundations of evaluation practice. This is how evaluation in the sense of evaluative research differs from the mere subjective judgement that the term 'evaluation' in its common sense may otherwise denote (Suchman 1967). Evaluation is first and foremost an applied research practice, and as such, it has borrowed a whole series of investigative techniques, both quantitative and qualitative, initially developed in basic research (e.g. questionnaires, quantitative analyses on databases, experimental methods, semi-structured interviews, observations, case studies, etc.). Beyond the techniques, the borrowing also concerns the methods of analysis and the conception of research designs. Despite this strong methodological link, evaluation does not boil down to a research practice (Wanzer 2021).This is suggested by the other two dimensions identified earlier (the concern for values and utility). In fact, the development of programme evaluation has given rise to a plurality of practices by a variety of public and private actors (public administration, consultants, NGOs, etc.), practices within which methodological issues are not necessarily central and where methodological rigour greatly varies. At the same time, the practice of evaluation has remained weakly and very unevenly institutionalised in the university (Cox 1990), where it suffers in particular from a frequent devaluation of applied research practices, suspicions of complacency towards commissioners, and difficulties linked to its interdisciplinary nature (see below) (Jacob 2008). Thus, although it has developed its journals and professional conferences, evaluation is still the subject of very few doctoral programmes and dedicated recruitments. Practised to varying degrees by different academic disciplines (public health, economics and development are now particularly involved), and sometimes described as 'transdisciplinary' in terms of its epistemological scope (Scriven 1993), evaluation is still far from being an academic discipline in the institutional sense of the term. From an epistemological point of view, this non- (or weak) disciplinarisation of evaluation is to be welcomed. The fact remains, however, that this leads to weaknesses. One of the consequences of this situation is a frequent lack of training for researchers in evaluation: particularly concerning the non-methodological dimensions of this practice (questions of values and utility), but also concerning certain approaches more specifically derived from evaluation practice. Indeed, while evaluation has largely borrowed from social science methods, it has also fostered a number of methodological innovations. For example, the use of experimental methods first took off in the social sciences in the context of evaluation, initially in education in the 1920s and then in social policy, health and other fields from the 1960s onwards (Campbell and Stanley 1963). The link with medicine brought about by the borrowing of the model of the clinical trial (the notions of 'trial' and 'treatment' having thus been transposed to evaluation) then favoured the transfer from the medical sciences to evaluation of another method, systematic literature reviews, which consists in adopting a systematic protocol to search for existing publications on (a) given evaluative question(s) and to draw up a synthesis of their contributions (Hong and Pluye 2018; Belaid and Ridde 2020). Without being the only place where it is deployed, programme evaluation has also made a major contribution to the development and theorising of mixed methods, which consist of articulating qualitative and quantitative techniques in the same research (Baïz and Revillard 2022; Greene, Benjamin and Goodyear 2001; Burch and Heinrich 2016; Mertens 2017). Similarly, because of its central concern with the use of knowledge, evaluation has been a privileged site for the development of participatory research and its theorisation (Brisolara 1998; Cousins and Whitmore 1998; Patton 2018). While these methods (experimental methods, systematic literature reviews, mixed methods, participatory research) are immediately applicable to fields other than programme evaluation, other methodological approaches and tools have been more specifically developed for this purpose. This is particularly the case of theory-based evaluation (Weiss 1997; Rogers and Weiss 2007), encompassing a variety of approaches (realist evaluation, contribution analysis, outcome harvesting, etc.) which will be described below (Pawson and Tilley 1997; Mayne 2012; Wilson-Grau 2018). Apart from a few disciplines in which they are more widespread, such as public health or development (Ridde and Dagenais 2009; Ridde et al. 2020), these approaches are still little known to researchers who have undergone traditional training in research methods, including those who may be involved in evaluation projects. A dialogue therefore needs to be renewed between evaluation and research: according to the reciprocal dynamic of the initial borrowing of research methods by evaluation, a greater diversity of basic research circles would now benefit from a better knowledge of the specific methods and approaches derived from the practice of evaluation. This is one of the vocations of LIEPP, which promotes a strengthening of exchanges between researchers and evaluation practitioners. Since 2020, LIEPP has been organising a monthly seminar on evaluation methods and approaches (METHEVAL), alternating presentations by researchers and practitioners, and bringing together a diverse audience. This is also one of the motivations behind the book Evaluation: Foundations, Controversies, Perspectives, published in 2021, which aimed in particular to make researchers aware of the non-methodological aspects of evaluation (Delahais et al. 2021). This publication completes the process by facilitating the appropriation of approaches developed in evaluation such as theory-based evaluation, realistic evaluation, contribution analysis and outcome harvesting. Conversely, LIEPP believes that evaluation would benefit from being more open to methodological tools more frequently used in basic research and with which it tends to be less familiar, particularly because of the targeting of questions at the scale of the intervention. In fact, evaluation classically takes as its object an intervention or a programme, usually on a local, regional or national scale, and within a sufficiently targeted questioning perimeter to allow conclusions to be drawn regarding the consequences of the intervention under study. By talking about policy evaluation rather than programme evaluation in the strict sense, our aim is to include the possibility of reflection on a more macro scale in both the geographical and temporal sense, by integrating reflections on the historicity of public policies, on the arrangement of different interventions in a broader policy context (a welfare regime, for example), and by relying more systematically on international comparative approaches. Evaluation, in other words, must be connected to policy analysis – an ambition already stated in the 1990s by the promoters of an "évaluation à la française" (Duran, Monnier, and Smith 1995; Duran, Erhel, and Gautié 2018). This is made possible, for example, by comparative historical analysis and macro-level comparisons presented in this book. Another important implication of programme evaluation is that the focus is on the intervention under study. By shifting the focus, many basic research practices can provide very useful insights in a more prospective way, helping to understand the social problems targeted by the interventions. All the thematic research conducted in the social sciences provides very useful insights for evaluation in this respect (Rossi, Lipsey, and Freeman 2004). Among the methods presented in this book, experimental approaches such as laboratory experimentation or testing, which are not necessarily focused on interventions as such, help to illustrate this more prospective contribution of research to evaluation. A dialogue between qualitative and quantitative approaches By borrowing its methods from the social sciences, policy evaluation has also inherited the associated methodological and epistemological controversies. Although there are many calls for reconciliation, although evaluation is more likely to emphasise its methodological pragmatism (the evaluative question guides the choice of methods), and although it has been a driving force in the development of mixed methods, in practice, in evaluation as in research, the dialogue between quantitative and qualitative traditions (especially in their epistemological dimension) is not always simple. Articulating different disciplinary and methodological approaches to evaluate public policies is the founding ambition of LIEPP. The difficulties of this dialogue, particularly on an epistemological level (opposition between positivism and constructivism), were identified at the creation of the laboratory (Wasmer and Musselin 2013). Over the years, LIEPP has worked to overcome these obstacles by organizing a more systematic dialogue between different methods and disciplines in order to enrich evaluation: through the development of six research groups co-led by researchers from different disciplines, through projects carried out by interdisciplinary teams, but also through the regular discussion of projects from one discipline or family of methods by specialists from other disciplines or methods. It is also through these exchanges that the need for didactic material to facilitate the understanding of quantitative methods by specialists in qualitative methods, and vice-versa, has emerged. This mutual understanding is becoming increasingly difficult in a context of growing technicisation of methods. This book responds to this need, drawing heavily on the group of researchers open to interdisciplinarity and to the dialogue between methods that has been built up at LIEPP over the years: among the 25 authors of this book, nine are affiliated to LIEPP and eight others have had the opportunity to present their research at seminars organised by LIEPP. This book has therefore been conceived as a means of encouraging a dialogue between methods, both within LIEPP and beyond. The aim is not necessarily to promote the development of mixed-methods research, although the strengths of such approaches are described (Part III). It is first of all to promote mutual understanding between the different methodological approaches, to ensure that practitioners of qualitative methods understand the complementary contribution of quantitative methods, their scope and their limits, and vice versa. In doing so, the approach also aims to foster greater reflexivity in each methodological practice, through a greater awareness of what one method is best suited for and the issues for which other methods are more relevant. While avoiding excessive technicality, the aim is to get to the heart of how each method works in order to understand concretely what it allows and what it does not allow. We are betting that this practical approach will help to overcome certain obstacles to dialogue between methods linked to major epistemological oppositions (positivism versus constructivism, for example) which are not necessarily central in everyday research practice. For students and non-academic audiences (particularly among policymakers or NGOs who may have recourse to programme evaluations), the aim is also to promote a more global understanding of the contributions and limitations of the various methods. Far from claiming to be exhaustive, the book aims to present some examples of three main families of methods or approaches: quantitative methods, qualitative methods, and mixed methods and cross-sectional approaches in evaluation. In what follows, we present the general organisation of the book and the different chapters, integrating them into a more global reflection on the distinction between quantitative and qualitative approaches. At a very general level, quantitative and qualitative methods are distinguished by the density and breadth of the type of information they produce: whereas quantitative methods can produce limited information on a large number of cases, qualitative methods provide denser, contextualised information on a limited number of cases. But beyond these descriptive characteristics, the two families of methods also tend to differ in their conception of causality. This is a central issue for policy evaluation which, without being restricted to this question, was founded on investigating the impact of public interventions: to what extent can a given change observed be attributed to the effect of a given intervention? – In other words, a causal question (can a cause-and-effect relationship be established between the intervention and the observed change?). To understand the complementary contributions of quantitative and qualitative methods for evaluation, it is therefore important to understand the different ways in which they tend to address this central question of causality. Quantitative methods Experimental and quasi-experimental quantitative methods are based on a counterfactual view of causality: to prove that A causes B, it must be shown that, all other things being equal, if A is absent, B is absent (Woodward 2003). Applied to the evaluation of policy impact, this logic invites us to prove that an intervention causes a given impact by showing that in the absence of this intervention, all other things being equal, this impact does not occur (Desplatz and Ferracci 2017). The whole difficulty then consists of approximating as best as possible these 'all other things being equal' situations: what would have happened in the absence of the intervention, all other characteristics of the situation being identical? It is this desire to compare situations with and without intervention 'all other things being equal' that gave rise to the development of experimental methods in evaluation (Campbell and Stanley 1963; Rossi, Lipsey, and Freeman 2004). Most experiments conducted in policy evaluation are field experiments, in the sense that they study the intervention in situation, as it is actually implemented. Randomised controlled trials (RCTs) (see Chapter 1) compare an experimental group (receiving the intervention) with a control group, aiming for equivalence of characteristics between the two groups by randomly assigning participants to one or the other group. This type of approach is particularly well suited to interventions that are otherwise referred to as 'experiments' in public policy (Devaux-Spatarakis 2014). These are interventions that public authorities launch in a limited number of territories or organisations to test their effects, thus allowing for the possibility of control groups. When this type of direct experimentation is not possible, evaluators can resort to several quasi-experimental methods, aiming to reconstitute comparison groups from already existing situations and data (thus without manipulating reality, unlike experimental protocols) (Fougère and Jacquemet 2019). The difference-in-differences method uses a time marker at which one of the two groups studied receives the intervention and the other does not, and measures the impact of the intervention by comparing the results before and after this time (see Chapter 2). Discontinuity regression (see Chapter 3) reconstructs a target group and a control group by comparing the situations on either side of an eligibility threshold set by the policy under study (e.g. eligibility for the intervention at a given age, income threshold, etc.). Finally, matching methods (see Chapter 4) consist of comparing the situations of beneficiaries of an intervention with those of non-beneficiaries with the most similar characteristics. In addition to these methods, which are based on real-life data, other quantitative impact assessment approaches are based on computer simulations or laboratory experiments. Microsimulation (see Chapter 5), the development of which has been facilitated by improvements in computing power, consists of estimating ex ante the expected impact of an intervention by taking into consideration a wide variety of data relating to the targeted individuals and simulating changes in their situation (e.g. ageing, changes in the labour market, fiscal policies, etc.). It also allows for a refined ex post analysis of the diversity of effects of a given policy on the targeted individuals. Policy evaluation can also rely on laboratory experiments (see Chapter 6), which make it possible to accurately measure the behaviour of individuals and, in particular, to uncover unconscious biases. Such analyses can, for example, be very useful in helping to design anti-discrimination policies, as part of an ex ante evaluation process. It is also in the context of reflection on these policies that testing methods (see Chapter 7) have been developed, making it possible to measure discrimination by sending fictitious applications in response to real offers (for example, job offers). But evaluation also seeks to measure the efficiency of interventions, beyond their impact. This implies comparing the results obtained with the cost of the policy under study and with those of alternative policies, in a cost-efficiency analysis approach (see Chapter 8). Qualitative methods While they are also compatible with counterfactual approaches, qualitative methods are more likely to support a generative or processual conception of causality (Maxwell 2004; 2012; Mohr 1999). Following this logic, causality is inferred, not from relations between variables, but from the analysis of the processes through which it operates. While the counterfactual approach establishes whether A causes B, the processual approach shows how (through what series of mechanisms) A causes B, through observing the empirical manifestations of these causal mechanisms that link A and B. In so doing, it goes beyond the behaviourist logic which, in counterfactual approaches, conceives the intervention according to a stimulus-response mechanism, the intervention itself then constituting a form of black box. Qualitative approaches break down the intervention into a series of processes that contribute to producing (or preventing) the desired result: this is the general principle of theory-based evaluations (presented in the third part of this book in Chapter 20 as they are also compatible with quantitative methods). This finer scale analysis is made possible by focusing on a limited number of cases, which are then studied in greater depth using different qualitative techniques. Particular attention is paid to the contexts, as well as to the mental processes and the logic of action of the people involved in the intervention (agents responsible for its implementation, target groups), in a comprehensive approach (Revillard 2018). Unlike quantitative methods, qualitative methods cannot measure the impact of a public policy; they can, however, explain it (and its variations according to context), but also answer other evaluative questions such as the relevance or coherence of interventions. Table 1 summarises these ideal-typical differences between quantitative and qualitative methods: it is important to specify that we are highlighting here the affinities of a given family of methods with a given approach to causality and a given consideration of processes and context, but this is an ideal-typical distinction which is far from exhausting the actual combinations in terms of methods and research designs. The most emblematic qualitative research technique is probably direct observation or ethnography, coming from anthropology, which consists of directly observing the social situation being studied in the field (see Chapter 9). A particularly engaging method, direct observation is very effective in uncovering all the intermediate policy processes that contribute to producing its effects, as well as in distancing official discourse through the direct observation of interactions. The semi-structured interview (see Chapter 10) is another widely used qualitative research technique, which consists of a verbal interaction solicited by the researcher with a research participant, based on a grid of questions used in a very flexible manner. The interview aims both to gather information and to understand the experience and worldview of the interviewee. This method can also be used in a more collective setting, in the form of focus groups (see Chapter 11) or group interviews (see Chapter 12). As Ana Manzano points out in her chapter on focus groups, the terminologies for these group interview practices vary. Our aim in publishing two chapters on these techniques is not to rigidify the distinction but to provide two complementary views on these frequently used methods. Although case studies (see Chapter 13) can use a variety of qualitative, quantitative and mixed methods, they are classically part of a qualitative research tradition because of their connection to anthropology. They allow interventions to be studied in context and are particularly suited to the analysis of complex interventions. Several case studies can be combined in the evaluation of the same policy; the way in which they are selected is then decisive. Process tracing (see Chapter 14), which relies mainly but not exclusively on qualitative enquiry techniques, focuses on the course of the intervention in a particular case, seeking to trace how certain actions led to others. The evaluator then acts as a detective looking for the "fingerprints" left by the mechanisms of change. The approach makes it possible to establish under what conditions, how and why an intervention works in a particular case. Finally, comparative historical analysis combines the two fundamental methodological tools of social science, comparison and history, to help explain large-scale social phenomena (see Chapter 15). It is particularly useful for reporting on the definition of public policies. Mixed methods and cross-cutting approaches in evaluation The third and final part of the book brings together a series of chapters on the articulation between qualitative and quantitative methods as well as on cross-cutting approaches that are compatible with a diversity of methods. Policy evaluation has played a driving role in the formalisation of the use of mixed methods, leading in particular to the distinction between different strategies for linking qualitative and quantitative methods (sequential exploratory, sequential explanatory or convergent design) (see Chapter 16). Even when the empirical investigation mobilises only one type of method, it benefits from being based on a systematic mixed methods literature review. While the practice of systematic literature reviews was initially developed to synthesise results from randomised controlled trials, this practice has diversified over the years to include other types of research (Hong and Pluye 2018). The particularity of systematic mixed methods literature reviews is that they include quantitative, qualitative and mixed studies, making it possible to answer a wider range of evaluative questions (see Chapter 17). Having set out this general framework on mixed methods and reviews, the following chapters present six cross-cutting approaches. The first two, macro-level comparisons and qualitative comparative analysis (QCA), tend to be drawn from basic research practices, while the other four (theory-based evaluation, realist evaluation, contribution analysis, outcome harvesting) are drawn from the field of evaluation. Macro-level comparisons (see Chapter 18) consist of exploiting variations and similarities between large entities of analysis (e.g. states or regions) for explanatory purposes: for example, to explain differences between large social policy models, or the influence of a particular family policy configuration on women's employment rate. Qualitative comparative analysis (QCA) is a mixed method which consists in translating qualitative data into a numerical format in order to systematically analyse which configurations of factors produce a given result (see Chapter 19). Based on an alternative, configurational conception of causality, it is useful for understanding why the same policy may lead to certain changes in some circumstances and not in others. Developed in response to the limitations of experimental and quasi-experimental approaches to understanding how an intervention produces its impacts, theory-based evaluation consists of opening the 'black box' of public policy by breaking down the different stages of the causal chain linking the intervention to its final results (see Chapter 20). The following chapters fall broadly within this family of evaluation approaches. Realist evaluation (see Chapter 21) conceives of public policies as interventions that produce their effects through mechanisms that are only triggered in specific contexts. By uncovering context-mechanism-outcomes (CMO) configurations, this approach makes it possible to establish for whom, how and under what circumstances an intervention works. Particularly suited to complex interventions, contribution analysis (see Chapter 22) involves the progressive formulation of 'contribution claims' in a process involving policy stakeholders, and then testing these claims systematically using a variety of methods. Outcome harvesting (see Chapter 23) starts from a broad understanding of observable changes, and then traces whether and how the intervention may have played a role in producing them. Finally, the last chapter is devoted to an innovative approach to evaluation, based on the concept of cultural safety initially developed in nursing science (see Chapter 24). Cultural safety aims to ensure that the evaluation takes place in a 'safe' manner for stakeholders, and in particular for the minority communities targeted by the intervention under study, i.e. that the evaluation process avoids reproducing mechanisms of domination (aggression, denial of identity, etc.) linked to structural inequalities. To this end, various participatory techniques are used at all stages of the evaluation. This chapter is thus an opportunity to emphasise the importance of participatory dynamics in evaluation, also highlighted in several other contributions. A didactic and illustrated presentation To facilitate reading and comparison between methods and approaches, each chapter is organised according to a common outline based on five main questions: 1) What does this method/approach consist of? 2) How is it useful for policy evaluation? 3) An example of the use of this method/approach; 4) What are the criteria for judging the quality of the use of this method/approach? 5) What are the strengths and limitations of this method/approach compared to others? The book is published directly in two languages (French and English) in order to facilitate its dissemination. The contributions were initially written in one or the other language according to the preference of the authors, then translated and revised (where possible) by them. A bilingual glossary is available below to facilitate the transition from one language to the other. The examples used cover a wide range of public policy areas, studied in a variety of contexts: pensions in Italy, weather and climate information in Senegal, minimum wage in New Jersey, reception in public services in France, child development in China, the fight against smoking among young people in the United Kingdom, health financing in Burkina Faso, the impact of a summer school on academic success in the United States, soft skills training in Belgium, the development of citizen participation to improve public services in the Dominican Republic, a nutrition project in Bangladesh, universal health coverage in six African countries, etc. The many examples presented in the chapters illustrate the diversity and current vitality of evaluation research practices. Far from claiming to be exhaustive, this publication is an initial summary of some of the most widely used methods. The collection is intended to be enriched by means of publications over time in the open access collection of LIEPP methods briefs. Cited references Alkin, Marvin. and Christie, Christina. 2012. 'An Evaluation Theory Tree'. In Evaluation Roots, edited by Marvin Alkin and Christina Christie. London: Sage. . Baïz, Adam. and Revillard, Anne. 2022. Comment Articuler Les Méthodes Qualitatives et Quantitatives Pour Évaluer L'impact Des Politiques Publiques? Paris: France Stratégie. . Belaid, Loubna. and Ridde, Valéry. 2020. 'Une Cartographie de Quelques Méthodes de Revues Systématiques'. Working Paper CEPED N°44. Brisolara, Sharon. 1998. 'The History of Participatory Evaluation and Current Debates in the Field'. New Directions for Evaluation, (80): 25–41. . Burch, Patricia. and Heinrich, Carolyn J.. 2016. Mixed Methods for Policy Research and Program Evaluation. Los Angeles: Sage. Campbell, Donald T.. and Stanley, Julian C.. 1963. 'Experimental and Quasi-Experimental Designs for Research'. In Handbook of Research on Teaching. Houghton Mifflin Company. Cousins, J. Bradley. and Whitmore, Elizabeth. 1998. 'Framing Participatory Evaluation'. New Directions for Evaluation, (80): 5–23. Cox, Gary. 1990. 'On the Demise of Academic Evaluation'. Evaluation and Program Planning, 13(4): 415–19. Delahais, Thomas. 2022. 'Le Choix Des Approches Évaluatives'. In L'évaluation En Contexte de Développement: Enjeux, Approches et Pratiques, edited by Linda Rey, Jean Serge Quesnel, and Vénétia Sauvain, 155–80. Montréal: JFP/ENAP. Delahais, Thomas, Devaux-Spatarakis, Agathe, Revillard, Anne, and Ridde, Valéry. eds. 2021. Evaluation: Fondements, Controverses, Perspectives. Québec: Éditions science et bien commun. . Desplatz, Rozenn. and Ferracci, Marc. 2017. Comment Évaluer l'impact Des Politiques Publiques? Un Guide à l'usage Des Décideurs et Praticiens. Paris: France Stratégie. Devaux-Spatarakis, Agathe. 2014. 'L'expérimentation "telle Qu'elle Se Fait": Leçons de Trois Expérimentations Par Assignation Aléatoire'. Formation Emploi, (126): 17–38. Duran, Patrice. and Erhel, Christine. and Gautié, Jérôme. 2018. 'L'évaluation des politiques publiques'. Idées Économiques et Sociales, 193(3): 4–5. Duran, Patrice. and Monnier, Eric. and Smith, Andy. 1995. 'Évaluation À La Française: Towards a New Relationship between Social Science and Public Action'. Evaluation, 1(1): 45–63. . Fougère, Denis. and Jacquemet, Nicolas. 2019. 'Causal Inference and Impact Evaluation'. Economie et Statistique, (510-511–512): 181–200. . Greene, Jennifer C.. and Lehn, Benjamin. and Goodyear, Leslie. 2001. 'The Merits of Mixing Methods in Evaluation'. Evaluation, 7(1): 25–44. Hong, Quan Nha. and Pluye, Pierre. 2018. 'Systematic Reviews: A Brief Historical Overview'. Education for Information, (34): 261–76. Jacob, Steve. 2008. 'Cross-Disciplinarization a New Talisman for Evaluation?' American Journal of Evaluation 19(2): 175–94. . Mathison, Sandra. 2005. Encyclopedia of Evaluation. London: Sage. Maxwell, Joseph A. 2004. 'Using Qualitative Methods for Causal Explanation'. Field Methods, 16(3): 243–64. . ———. 2012. A Realist Approach for Qualitative Research. London: Sage. Mayne, John. 2012. 'Contribution Analysis: Coming of Age?' Evaluation, 18(3): 270–80. . Mertens, Donna M. 2017. Mixed Methods Design in Evaluation. Thousand Oaks: Sage. Evaluation in Practice Series. . Mohr, Lawrence B. 1999. 'The Qualitative Method of Impact Analysis'. American Journal of Evaluation, 20(1): 69–84. . Newcomer, Kathryn E.. and Hatry, Harry P.. and Wholey, Joseph S.. 2015. Handbook of Practical Program Evaluation. Hoboken: Wiley. Patton, Michael Q. 1997. Utilization-Focused Evaluation. 3rd ed. Thousand Oaks: Sage. ———. 2015. Qualitative Research and Evaluation Methods: Integrating Theory and Practice. London: Sage. ———. 2018. Utilization-Focused Evaluation. London: Sage. Pawson, Ray. and Tilley, Nicholas. 1997. Realistic Evaluation. London: Sage. Revillard, Anne. 2018. Quelle place pour les méthodes qualitatives dans l'évaluation des politiques publiques? Paris: LIEPP Working Paper n°81. Ridde, Valéry. and Dagenais, Christian. 2009. Approches et Pratiques En Évaluation de Programme. Montréal: Presses de l'Université de Montréal. Ridde, Valéry. and Dagenais, Christian (eds). 2020. Évaluation Des Interventions de Santé Mondiale: Méthodes Avancées. Québec: Éditions science et bien commun. Rogers, Patricia J.. and Weiss, Carol H.. 2007. 'Theory-Based Evaluation: Reflections Ten Years on: Theory-Based Evaluation: Past, Present, and Future'. New Directions for Evaluation, (114): 63–81. . Rossi, Peter H.. and Mark W. Lipsey. and Howard E. Freeman. 2004. Evaluation: A Systematic Approach. London: Sage. Scriven, Michael. 1993. 'Hard-Won Lessons in Program Evaluation.' New Directions for Program Evaluation, (58): 5–48. Suchman, Edward A. 1967. Evaluative Research. Principles and Practice in Public Service and Social Action Programs. New York: Russell Sage Foundation. Wanzer, Dana L. 2021. 'What Is Evaluation? Perspectives of How Evaluation Differs (or Not) From Research'. American Journal of Evaluation, 42(1): 28–46. . Wasmer, Etienne. and Musselin, Christine. 2013. Évaluation Des Politiques Publiques: Faut-Il de l'interdisciplinarité? Paris: LIEPP Methodological discussion paper n°2. Weiss, Carol H. 1997. 'Theory-Based Evaluation: Past, Present, and Future'. New Directions for Evaluation (76): 41–55. . ———. 1998. Evaluation: Methods for Studying Programs and Policies. Upper Saddle River, NJ: Prentice-Hall. Wilson-Grau, Ricardo. 2018. Outcome Harvesting: Principles, Steps, and Evaluation Applications. IAP. Woodward, James. 2003. Making Things Happen a Theory of Causal Explanation. Oxford Studies in Philosophy of Science. New York: Oxford University Press. What is Policy Evaluation? Policy evaluation can be better defined as a process by which general judgments about quality, goal attainment, program effectiveness, impact, and costs can be determined (Theodoulou and Kofinis, 2004, p 192) Once public policy has been operationalized through the formal adoption of laws, rules, or regulations, and the bureaucracy has taken action to implement the policy, some form of evaluation needs to be accomplished to determine if the policy has achieved the desired outcome or impact. Public policy represents the expenditure of limited public resources and or restrictions on certain types of individual or organizational behavior. Consequently, the public has a right to expect that their government officials are accountable for the validity, efficiency, and effectiveness of those policies. Policy evaluation is therefore an absolutely critical stage in the policy process whereby ìwe can determine whether a policyís effects are intended or unintended and whether the results are positive or negative for the target population and society as a whole (Theodoulou and Kofinis, 2004, p. 191).î In essence, policy evaluation is the process used to determine what the consequences of public policy are and what has and has not been achieved. What is Policy Evaluation? Elected officials, policy makers, community leaders, bureaucrats, and the public want to know what policies work and what policies don't. The purpose of evaluation is to determine whether an implemented program is doing what it is supposed to. Through evaluation, we can determine whether a policy's effects are intended or unintended and whether the results are positive or negative for the target population and society as a whole. (Theodoulou and Kofinis, 2004, p. 191) At first glance, policy evaluation appears to be a straightforward concept however a closer inspection of the process reveals that policy evaluation can be equally as political and divisive as any other stage of the policy process. Policy evaluation provides additional opportunities for the myriad political interest groups and policy actors to attempt to influence the life of a specific policy. Favorable evaluations of the impact of a given policy will tend to perpetuate the implementation and life-cycle while unfavorable evaluations may give rise to change or possibly policy termination. Depending on the proclivity of any interest group or policy actor, the perception of how well a policy or program is performing or being implemented can have far-reaching impact. The retrospective analysis of any public policy or government action is bounded by a number of ìreal-world constraints, such as time, budget, ethical considerations, and policy restrictions (Theodoulou and Kofinis, 2004, p. 193),î as well as political ideologies, values, experiences, measurement instruments, goal clarity, and institutional biases. The key to understanding and interpreting the results of any policy evaluation is that some degree of bias is inherent in the process. However, this shortcoming should not prevent efforts to produce fair and unbiased policy evaluation products, at least as much as possible. The objective of policy evaluation is to discover policy flaws and to attempt to correct them given all of the limitations incumbent in the overarching policy process. ìIn its simplest form, evaluating a public program involves cataloging the goals of the program, measuring the degree to which the goals have been achieved, and, perhaps, suggesting changes that might bring the performance of the organization more in line with the stated purposes of the program (Peters, 2007, p. 163).î What is Policy Evaluation? The consequences of such policy programs are determined by describing their impacts, or by looking at whether they have succeeded or failed according to a set of established standards. Several evaluation perspectives are: Evaluation is the assessment of whether a set of activities implemented under a specific policy has achieved a given set of objectives. Evaluation is the effort that renders a judgment about program quality. Evaluation is information gathering for the purposes of making decisions about the future of the program. Evaluation is the use of scientific methods to determine how successful implementation and its outcomes have been. (Theodoulou and Kofinis, 2004, p. 192) Policy evaluation plays a crucial role in determining whether public policies are achieving their intended goals and making a meaningful impact on society. The evaluation process helps policymakers understand what works, what doesn't, and how policies can be improved for better outcomes. Various entities perform policy evaluations, each bringing unique perspectives and methodologies to the assessment process. Table of Contents Special agencies dedicated to policy evaluation represent one of the most formal and structured approaches to assessing public policies. These agencies are typically established with the specific mandate of conducting professional evaluations across various government programs and initiatives. Characteristics of special agency evaluations Special evaluation agencies are characterized by their professional expertise, independence, and methodological rigor. These organizations employ trained evaluators with specialized knowledge in research methodologies, data analysis, and program assessment techniques. In India, organizations such as NITI Aayog, Development Monitoring and Evaluation Office (DMEO), and Program Evaluation Organization (PEO) serve as dedicated evaluation bodies that systematically assess the effectiveness and efficiency of government policies and programs. These agencies employ a variety of evaluation frameworks and methodologies tailored to the specific policy context. Advantages of special agency evaluations Professional expertise: Special agencies employ trained evaluators with expertise in evaluation methodologies and analytical techniques. Independence and objectivity: Being separate from implementing agencies, these evaluators can provide unbiased assessments without being influenced by operational considerations. Comprehensive approach: These agencies can conduct thorough evaluations that examine multiple dimensions of policy performance, including efficiency, effectiveness, relevance, and sustainability. Methodological rigor: Special agencies typically employ sophisticated research designs and analytical methods that enhance the credibility of their findings. Limitations of special agency evaluations Distance from implementation: Evaluators may lack the intimate knowledge of ground realities that implementing officials possess. Resource constraints: Comprehensive evaluations require significant resources and time, which may limit the scope and frequency of evaluations. Potential for bureaucratic distance: Findings may not always be effectively communicated to or utilized by policy implementers and decision-makers. Operating staff as policy evaluators Operating staff—the civil servants and government employees directly involved in implementing policies—represent another important source of policy evaluation. Their day-to-day involvement with policy implementation gives them unique insights into how policies function on the ground. Operating staff evaluations are characterized by their practical orientation and grounding in implementation realities. These evaluations often focus on operational aspects of policies, including implementation challenges, procedural bottlenecks, and resource adequacy. In India, operating staff at various levels of government—from central ministries to district and block-level implementers—regularly assess policy implementation through internal reviews, monitoring exercises, and feedback mechanisms. Their evaluations tend to be more continuous and integrated into the policy implementation process. Advantages of operating staff evaluations Detailed knowledge: Implementing officials possess intimate knowledge of policy operations, target populations, and contextual factors that influence policy outcomes. Operational focus: Their evaluations tend to highlight practical implementation issues and constraints that might be overlooked in more theoretical assessments. Real-time insights: Operating staff can provide continuous feedback on policy performance rather than point-in-time assessments. Solution-oriented: Their evaluations often include practical suggestions for improving policy implementation based on frontline experiences. Limitations of operating staff evaluations Lack of objectivity: Implementing officials may have vested interests in portraying policies favorably, potentially compromising the objectivity of their assessments. Methodological limitations: Operating staff may lack specialized evaluation expertise and rely more on anecdotal evidence than systematic data collection and analysis. Narrow focus: These evaluations may emphasize immediate operational concerns over broader policy impacts and outcomes. Limited perspective: Staff may not see the bigger picture or understand how their program fits into broader policy objectives. Legislative committees as policy evaluators Legislative bodies play a crucial oversight role in democratic governance, and their evaluative function is exercised primarily through legislative committees. These committees represent the interests of citizens and serve as important accountability mechanisms for executive actions. Parliamentary oversight in policy evaluation In India, parliamentary committees such as the Public Accounts Committee, Estimates Committee, and various Standing Committees regularly examine policy implementation and outcomes. These committees hold hearings, review reports, summon officials, and issue recommendations on policy matters. Legislative committee evaluations typically focus on policy compliance with legislative intent, budgetary appropriations, and broader public interest considerations. They serve as important forums for holding implementing agencies accountable and for amplifying citizen concerns about policy impacts. Advantages of legislative committee evaluations Democratic legitimacy: As representatives of citizens, legislators bring democratic legitimacy to the evaluation process and can voice public concerns about policies. Broad perspective: Legislative committees often examine policies from multiple angles, including legal, fiscal, social, and political dimensions. Public transparency: Committee proceedings and reports are typically public, enhancing transparency in policy assessment. Leverage for change: Legislative committees can use budgetary powers and public pressure to incentivize policy improvements based on evaluation findings. Limitations of legislative committee evaluations Political considerations: Evaluations may be influenced by political affiliations and partisan interests rather than objective assessment criteria. Technical limitations: Legislators may lack specialized expertise in policy domains or evaluation methodologies. Episodic attention:

Legislative oversight tends to be intermittent rather than continuous, potentially missing ongoing implementation issues. Resource constraints: Committees often have limited staff and resources to conduct comprehensive evaluations across multiple policy areas. The Office of the Comptroller and Auditor General (CAG) as policy evaluator The Office of the Comptroller and Auditor General represents a constitutional authority with a mandate to examine public expenditures and ensure financial propriety in government operations. While traditionally focused on financial audits, the CAG has increasingly engaged in performance audits that evaluate policy outcomes and effectiveness. CAG's expanding evaluative role In India, the CAG conducts performance audits of major government schemes and programs, examining not just financial compliance but also economy, efficiency, and effectiveness in policy implementation. The CAG's reports often highlight implementation gaps, resource utilization issues, and shortfalls in achieving policy objectives. The CAG's evaluations are characterized by their constitutional authority, methodological rigor, and focus on accountability. These evaluations typically employ standardized audit procedures, extensive data collection, and comprehensive analysis of policy documentation and implementation records. Advantages of CAG evaluations Constitutional authority: The CAG's constitutional status provides legitimacy and influence to its evaluation findings. Independence: As an independent institution, the CAG can provide objective assessments without executive influence. Methodological rigor: CAG audits follow standardized methodologies and procedures that enhance the credibility of findings. Comprehensive coverage: The CAG can examine policies across multiple departments and jurisdictions, providing a more holistic assessment. Limitations of CAG evaluations Post-facto assessment: CAG evaluations typically occur after policy implementation, limiting their ability to influence real-time adjustments. Emphasis on compliance: Despite expanding into performance audits, CAG evaluations may still emphasize financial and procedural compliance over broader policy impacts. Time lag: The time between policy implementation and CAG evaluation can be substantial, potentially reducing the relevance of findings. Limited follow-up mechanisms: The CAG can identify issues but relies on executive and legislative action to implement recommendations. Statutory commissions as policy evaluators Various statutory commissions established to protect and promote specific public interests also play important evaluative roles in the policy process. These commissions focus on how policies affect particular population groups or issue areas under their mandate. Specialized evaluation focus In India, statutory bodies such as the National Human Rights Commission, National Commission for Women, National Commission for Scheduled Castes, and National Commission for Scheduled Tribes regularly evaluate policies affecting their respective domains. These evaluations typically examine whether policies adequately address the needs and rights of specific populations. Commission evaluations often employ rights-based frameworks and assess policies against constitutional provisions, legal mandates, and international commitments. They frequently incorporate participatory methods that elevate the voices of affected populations in the evaluation process. Advantages of statutory commission evaluations Specialized expertise: Commissions possess domain-specific knowledge and understanding of the issues affecting particular populations. Rights perspective: These evaluations typically emphasize rights and justice considerations that might be overlooked in more technocratic assessments. Representation of marginalized voices: Commissions often amplify the perspectives of groups whose views might otherwise be underrepresented in policy evaluations. Statutory authority: Their legal mandate gives weight to their evaluation findings and recommendations. Limitations of statutory commission evaluations Narrow focus: Commission evaluations may emphasize specific aspects of policies relevant to their mandate rather than comprehensive assessment. Resource constraints: Many commissions operate with limited resources, restricting the scope and depth of their evaluative activities. Limited enforcement powers: While commissions can make recommendations, they typically lack direct authority to enforce policy changes. Potential for isolation: Commission evaluations may not always be effectively integrated with other evaluation streams or mainstream policy processes. Integrated approaches to policy evaluation The various forms of policy evaluation are not mutually exclusive but can complement each other in creating a more comprehensive understanding of policy performance. An integrated approach to policy evaluation leverages the strengths of different evaluation forms while mitigating their respective limitations. Effective policy evaluation systems often incorporate multiple evaluation streams, with different entities conducting assessments according to their comparative advantages. For example, operating staff might provide continuous feedback on implementation issues, special agencies might conduct periodic impact assessments, and the CAG might examine resource utilization and compliance aspects. The integration of these various evaluation forms can provide a more holistic picture of policy performance and contribute to more informed decision-making. Such integration requires coordination mechanisms, shared evaluation standards, and platforms for synthesizing diverse evaluation findings. In India, the establishment of evaluation frameworks that incorporate inputs from various evaluative bodies represents an important step toward more comprehensive and useful policy assessments. The growing emphasis on evidence-based policymaking further underscores the importance of robust and multi-faceted evaluation approaches. What do you think? How can we better integrate these different forms of policy evaluation to create more comprehensive assessments? Do you believe that certain forms of policy evaluation are more credible or valuable than others in the Indian context?