Evaluation in Urban Planning: Advances and Prospects

Vitor Oliveira¹ and Paulo Pinho¹

Abstract
This article provides an overview on the current debate on evaluation in urban planning. An initial evaluation state of the art is presented in three parts: the evolution of evaluation theory and methods, the contemporary planning debate around different perspectives and paradigms, and the nature and extent of evaluation practice in planning. The second part of the article focuses on the growing emphasis on urban form issues in different planning systems. This comprehensive literature review provides the background to support the authors’ proposal for a set of general principles to evaluate the implementation of urban plans.

Keywords
planning evaluation, urban form, plan implementation, ongoing assessment methodologies

In the beginning of the second half of the twentieth century, planning became progressively associated with the rational-comprehensive model. One of the distinctive features of this then new way of thinking was the integration of evaluation in the plan making process. In a specific planning situation, the decision maker would consider all possible courses of action, according to a number of established ends, identify and assess all the consequences following from the adoption of each course of action, and then select the most preferable alternative. Another feature, distinguishing the rational paradigm from the classical paradigm associated to Patrick Geddes’s survey-analysis-plan, was the new way of considering the physical dimension of the city. While in the classical paradigm the emphasis was on the city—the survey works, the analysis of collected data, and the preparation of a plan intended to control the future of the city—in the rational paradigm the focus was on the process and the method, leading to an unintended devaluation of the city as the planning object.

This article focuses on the evaluation of urban planning, providing a review of the state of the art and a number of recommendations to improve its practice. When a growing number of studies and advances in planning evaluation are being disseminated, a comprehensive and up-to-date review of this field seems particularly relevant. An underlying concern throughout the whole article is the consideration of urban form as a fundamental theme in planning and in planning evaluation.

The first and the second parts of the article contain a literature review. The review on planning evaluation is structured in three parts: an analysis of the evolution of evaluation theory and methods covering the second half of the twentieth century, an overview of the contemporary debate on evaluation in urban and regional planning, and a first reflection on evaluation in planning practice.

Closely following Alexander (2006c) and Khakee (2003)’s approaches, we analyze the evolution of evaluation throughout the past fifty years from three different perspectives: a policy program perspective, a planning theory perspective, and a welfare economics perspective. The identification and the characterization of the first and of the second traditions highlight the development of generations of evaluation (Guba and Lincoln 1989) and the tensions between different planning paradigms. The third tradition makes evident the importance of evaluation methods and of some classification schemes such as the model presented by Söderbaum (1998), based on the degree of aggregation of each method.

A number of issues from the contemporary debate are subsequently approached: (1) the need for evaluation and its integration in the planning process, (2) the timing of the evaluation exercise, (3) the different conceptions of success in plan implementation, (4) the necessary adjustments between the evaluation methodology and the specific plan concept, (5) the evaluation questions, the criteria, and the indicators, and finally (6) the presentation of the evaluation results and their use by decision makers.

The analysis of planning evaluation practice starts with a reflection on the existing gap between evaluation theory and practice. Beside critical situations, as the quasi absence of

¹University of Oporto, Porto, Portugal

Corresponding Author:
Vitor Oliveira, University of Oporto, Faculty of Engineering, Porto, 4200-465, Portugal
Email: vitorm@fe.up.pt
evaluation practice in some countries, one of the most evident aspects of this gap is the dominance of quantitative methods in practice, in contrast to the research field that seems to prefer qualitative approaches. After this analysis, five evaluation methodologies are presented.

The second part of the literature review focuses on the importance of urban form within planning throughout the twentieth century. One of the most significant changes identified has been the progressive devaluation of the plan as the main product of planning activity and as the basic support mechanism to control the form of the city. This devaluation, which has steadily occurred throughout the second half of the twentieth century, has been particularly evident in planning theory writings and not so much in planning practice. Indeed, the focus of planning research has progressively moved toward the planning process, decision making, the planning discourse, and communicative practices. Nevertheless, and somehow expressing the cyclical dimension of the debate in any intellectual activity, the physical form of the city became again a central topic of debate in the past decade.

In the final part of the article, a set of general principles for plan evaluation is proposed. The definition of these guiding principles reflects the main theoretical contributions on the nature of the relationships between evaluation and planning, the much-needed bridges between planning theory and planning practice, the most adequate timings for the evaluation exercises, and, finally, the emphasis on the products, the processes, and the results.

I. Evaluation in Planning

I.1. The Evolution of Evaluation Theory and Methods

The policy program perspective is the first to be approached. In the absence of a general and consensual vision on the evolution of evaluation theory, many authors converge on the acknowledgment of a shift from a positivist paradigm to a constructivist paradigm. Some describe this evolution in four evaluation generations (Guba and Lincoln 1989; Khakee 1998, 2003). According to Guba and Lincoln (1989) the first generation of evaluation exercises was intended to measure individual attributes, the second generation was geared toward the description of programs and objectives, the third was dominated by judgments on the contextual values of the object, and the fourth was centered around the negotiation of claims, concerns, and issues presented by the different stakeholders (see Table 1). In this process, the role of the evaluator has also evolved, from an initial position where he or she mainly was a technician to later becoming a describer, then a judge, and finally a mediator. The fourth generation went beyond the strict scientific dimension, including the human, the political, the social, the cultural, and the contextual dimensions. These American authors identify two major foundations for their contribution, a responsive focusing and a constructivist methodology. In recent years, several authors have criticized this simplistic vision of the three positivist generations, adopting more consensual positions (Patton 2002; Knaap 2004) or moving away from the constructivist paradigm (Pawson and Tilley 1997).

The planning theory perspective is the second to be approached. Khakee (1998) argues that, from both a theoretical and a practical point of view, planning and evaluation are inseparable concepts. Assuming that a particular planning concept presupposes a particular evaluation type, the changes in planning theory affect the functions and the major characteristics of evaluation. Khakee defends a shift of paradigms from rational planning to communicative planning. Supported on the work of Innes (1995), the Swedish author analyzes eight theoretical positions developed throughout the past fifty years: rational-comprehensive planning, incremental planning, advocacy planning, implementation-oriented planning, strategic planning, transactive planning, negotiated planning, and communicative planning. Khakee argues that while some of these (from the second to the fifth) were developed as a response to the rational planning model, others (the sixth and the seventh) were developed as new ideas leading to communicative planning theory. This change of paradigms is also defended by Forrester (1989), Healey (1996), Innes (1995), and Sager (1994). Despite some
differences, also Voogd (1998) argues that as society is moving from a representative democracy to a participatory democracy, so traditional evaluation methods will increasingly have a limited use in physical planning.

On the contrary, authors such as Alexander (1998b, 2000) and Lichfield (1998, 2001a) defend the use of integrated approaches, in which rational planning sustains an important role. Mandelbaum (1979) asks whether any paradigm can answer all the normative questions of planners, given the complexity of planning, and concludes that a complete general theory of planning is impossible. Alexander (1998b, 2000) proposes a contingent framework integrating four different views of planning: rational planning, communicative practice, coordinative planning, and frame setting. In this framework, the four planning models are complementary, not conflicting. Each model involves different kinds of actors, with different activities, in different stages of the planning process. Consequently, Alexander (1998b, 2000) argues that there is no need to substitute the rational model because in some situations it is still the most adequate model. Also Faludi (2000, 2006) distinguishes between situations where planning is a technical exercise—and the rational model is the most adequate—and other situations where planning is a learning process and new approaches are needed. Lichfield (2001a) highlights the integrating capacity of the community impact evaluation, a method presented in Lichfield (1996). Although agreeing with the need for alternative approaches, Lichfield (2001a) defends that these should be integrated within a more complex and diverse rational model. He recognizes the changes that have been taking place in the evaluation and planning models required by a changing society. Nevertheless, and contrary to Alexander (1998b, 2000), this author seeks the integration of planning and evaluation within his model.

Rationality is another issue in this debate. Alexander (2000, 2006b) and Lichfield (1998) argue that rationality will always be associated with planning and with evaluation. The former describes the evolution of evaluation theories and methods according to the different underlying types of rationality: instrumental, substantive, bounded, strategic, and communicative. This author argues that discussion should shift from questioning the link between rationality and planning to asking what kinds of planning and rationality particular cases, situations, or contexts demand (Alexander 2000, 242).

The welfare economics perspective is the third to be approached. A preliminary introduction on five of the most popular evaluation methods developed in the past five decades must be done. Cost–benefit analysis (CBA) was initially conceived to evaluate American federal projects on water resources and U.K. investments in transports. Later this methodology was extended to the appraisal of several kinds of public actions and projects. It is probably the evaluation method most widely used. The technique, although complex, uses one main criterion—the willingness to pay—and follows a simple principle—the association of a monetary value to each identified effect of a project or action. To determine if the benefits, or the desired effects of a project, are larger than the costs, or the undesired effects, a number of monetary classifications are summed up in a grand index. McAllister (1982) synthesizes the main strengths of CBA as follows: it is based on an established theory of value, it attempts to reflect the values of all people, it uses understandable impact categories and measurement units, and it has a valuable and extensive body of literature on its applications.

Planning balance sheet analysis (PBSA) was presented by Lichfield in the 1950s and, soon after, applied by this author in a number of plans in Britain. It is an adaptation of CBA to urban and regional planning, sharing the basic theory and techniques of the former. PBSA goes beyond CBA in two aspects at least: it integrates nonquantifiable impacts, introducing symbols in the appraisal tables alongside with monetary impacts, and it records detailed information on costs and benefits, distinguishing how the different social groups and stakeholders will be affected by the proposed plan under analysis. More recently, Lichfield (1996) presented the community impact evaluation—a natural departure from PBSA—with similar foundations on CBA and on the impact evaluation tradition.

Hill (1968) presented the goals–achievement matrix (GAM) in his PhD thesis in 1966 and synthesized in Hill (1968), as an attempt to eliminate the weaknesses of the existing evaluation methods, particularly CBA and PBSA. This method has been used in the evaluation of urban plans in Great Britain. The main characteristics of GAM are the organization of effects according to the goals and to the different parties involved and the incorporation of nonmonetary effects in the grand index. The main stages of this method are the definition of goals and objectives in operational terms to enable the measurement of their achievement, the attribution of value weights to the goals, reflecting their importance, the identification of goals achievement to different groups, the sum of achievement levels in a grand index, and finally the adjustment of this index considering equity issues.

Multicriteria analysis (MA) emerged in the 1960s in France. The electrotechniques and concordance-techniques soon became dominant among the new evaluation methodologies. The large number of MA methods that emerged in the subsequent decade is classified by Voogd (1983) and Nijkamp, Rietveld, and Voogd (1990). The former uses three main characteristics for this classification—the way space is included in an evaluation, the nature of the data employed, and the measurement scale of the information. Nijkamp, Rietveld, and Voogd classify these methods according to the number of choice possibilities, and to the type of information utilized. In general, MA methods adopt the form of a matrix with at least two dimensions, one expressing the different project alternatives and the other expressing the objectives.
and the evaluation criteria. The relative importance of the different criteria is reflected by an appropriate set of priorities or weights. As Lichfield’s method is inseparable from CBA, contemporary MA is also strongly related to GAM.

Environmental impact assessment (EIA) was introduced in the American federal system back in 1970 to assess the environmental impacts of major federal projects. Nowadays, it constitutes one of the main policy instruments of most national and regional environmental administrations. The main objective of EIA is to produce better decisions, informing politicians and planners on the environmental consequences of their proposals and defining a number of measures to eliminate or minimize the forecasted impacts. The EIA process includes the preparation of an environmental impact study, a reviewing procedure, public participation, and often-times an ex post evaluation (Partidário and Pinho 2000). Strategic environmental assessment is another instrument of impact evaluation, geared toward the evaluation of policies, plans, and programs. It works with strategies instead of specific development proposals, operating on larger temporal and geographical scales, with growing uncertainty levels and thus requiring greater flexibility (Partidário 2003).

After this introduction on evaluation methods, a classification scheme is presented. Söderbaum (1998) uses the degree of aggregation to identify three different groups: highly aggregated methods, intermediate methods, and highly disaggregated methods. The highly aggregated methods intend to sum all impacts in terms of a single value. This implies the existence of a consensus in society about specific valuation rules. CBA is a clear example of this group of methods—the focus is on the quantitative ratio of benefits and costs. It is essentially a monetary method even when nonmonetary impacts are considered. Intermediate methods also use a single quantitative indicator to express the overall utility of an alternative, but in this case the indicator has a composite nature reflecting different dimensions. According to Khakee (2003) these methods have been used in recent years but they have been increasingly criticized because they do not pay sufficient attention to the conflicting values of individuals. PBSA and some types of multicriteria evaluation can be classified as intermediate methods. Highly disaggregated methods are intrinsically multidimensional and do not intend to show the overall value of the plan. On the contrary, assessment and display of different impacts intend to stimulate interactive discourses and consensus building. The design of these methods adapts to changing contexts. Not only results but also the way of arriving at them are important. These methods combine inductive and deductive analysis and make use of quantitative and qualitative information (Khakee 2003). This third set of methods includes EIA and the so-called positional analysis presented by Söderbaum (1998). Khakee (2003) identifies a link among this third set of methods, the communicative planning paradigm, and the fourth generation of evaluation (Guba and Lincoln 1989).

1.2. The Contemporary Debate on Planning Evaluation

Evaluation in the planning process. The evaluation of planning practice is a complex but most necessary exercise (Alexander 2006d; Alexander and Faludi 1989; Baer 1997; Brody and Highfield 2005; Brody, Highfield, and Thornton 2006; Talen 1997; Laurian, Day, Backhurst, et al. 2004). If planning intends to have any credibility as a discipline or as a profession, it should be possible, through a systematic assessment, to have a real judgment of planning effectiveness (Alexander and Faludi 1989). “Good” planning or “good” plans should be distinguishable from “bad” planning and “bad” plans (Alexander 2002; Alexander and Faludi 1989, 127; Baer 1997). The difficulties, the uncertainties, and the complexity of planning evaluation offer a context of limited rationality, in which evaluation conclusions have to be carefully bounded.

Although planning and evaluation should be two inseparable concepts, there is currently, as indeed three decades ago when Lichfield, Kettle, and Whitbread (1975) raised this question, a deficient integration of evaluation in the planning process. In a recent review of his seminal study from the seventies, Lichfield (2001b) raises some key issues: the evaluator should be, right from the beginning, a member of the planning team; the evaluation processes and criteria should be decided together with that team; the evaluation criteria should also be the design criteria; and the required information for the evaluation exercise should be decided in the beginning of the process.

The position of Lichfield (2001b) about the relationships between the evaluator and the planning team is not consensual in literature. Undoubtedly, there are advantages, but also disadvantages, in the integration of the evaluator in the planning team. The main advantages of an evaluation are fourfold: a greater knowledge on the specific institutional context, a greater probability of adopting the final recommendations, a reduced possibility to look at the assessment exercise as a threat to the institution, and eventually a reduced use of financial resources. The major disadvantage is the tendency to avoid negative conclusions and to accept the conventional line of thought. The main advantages of an external evaluation may be a greater objectivity in the evaluation and the facility of external hiring during short periods of time. The most important disadvantages mirror the main advantages of an internal exercise. This reflection should also consider two different possibilities, a mixed evaluation and an internal assessment prepared by an evaluator independent from the planning team.

The timing of the evaluation. Despite the specific characteristics of each country’s legal framework, the planning process is frequently oriented to the preparation, implementation, and revision of its main product, the plans (Lichfield and Prat 1998). Each of these stages corresponds, or should
correspond, to a particular stage in the evaluation process. *Ex ante* evaluation occurs in the beginning of the planning process and promotes the comparison of possible alternatives to choose the best solution for further development. *Ongoing* evaluation takes place during plan implementation, and its conclusions can lead to shifts in the planning process. Focusing on the plan results and on the use of resources, this kind of assessment requires a set of information that should be provided by an adequate data system. *Ex post* evaluation occurs at the end of the plan implementation process and focuses on the impacts of the plan. This type of evaluation reviews the whole process of preparation and implementation of the plan and formulates a judgment about its success.

The literature on planning evaluation consensually sustains the view that the study of the ongoing and *ex post* dimensions has a rather reduced expression, when compared to the analysis of the *ex ante* dimension (Berke et al. 2006; Brody and Highfield 2005; Brody et al. 2006b; Laurian, Day, Backhurst, et al. 2004; Lichfield 1996, 2001b, 2003). Lichfield (2001b) analyzes the recent development of evaluation in planning in relation to the evaluation of programs. In the former *ongoing* and *ex post* evaluation have a marginal role, whereas in the latter *ex ante* evaluation is usually devaluated because of the alleged difficulties of social sciences in providing reliable forecast. Accordingly, this British author challenges academics and professionals in both fields to compare their works and methodologies.

**Conceptions of success in plan implementation.** In the beginning of this article, five evaluation methods normally used during plan preparation were briefly presented. Planning evaluation methods used in *ex ante* evaluation exercises are based on *utilitarian* or *modified utilitarian* principles (Alexander 2006a). As a whole, these are distinct from the approaches presented in the next paragraphs covering the so-called *conformance-based* and *performance-based* evaluation.

Conformance-based evaluation means judging the success or failure of planning using one or two criteria—the conformance degree between the outcomes on the ground and the plan proposals and the promotion of planning goals and objectives through the available implementation instruments (Alexander 2006a). This approach has been developed, among others, by Alterman and Hill (1978), Baer (1997), Brody and Highfield (2005), Brody et al. (2006b), Burby (2003), Calkins (1979), Laurian, Day, Backhurst, et al. (2004), Laurian, Day, Berke, et al. (2004), and Talen (1996a, 1996b, 1997; see Table 2).

One of the first contributions to this approach was Alterman and Hill (1978)'s research work in Israel, measuring the conformance degree between the land use plan proposals and the urban development patterns. Also at the end of the 1970s, Calkins (1979) presented the *planning monitor*, a mechanism to measure the achievement of plan objectives and to explain eventual differences between planning and urban development. Influenced by these studies, Talen (1996a) analyzes the distribution of public facilities in a particular American city. In this article, as well as in a later article (Talen 1997), the author presents some proposals to improve this evaluation approach: planning literature should focus not only on planning failures but also on planning successes, a substantive and object-oriented view of planning should be promoted, and planners should incorporate an evaluation mechanism in each plan they prepare, establishing how to measure the goals achievement. The Plan Implementation Evaluation (PIE) is presented by Laurian, Day, Berke, et al. (2004) as an answer to Talen (1997)'s challenge to develop a methodology to assess plan implementation. These authors

### Table 2. Ongoing and Ex Post Evaluation Approaches

<table>
<thead>
<tr>
<th>Conformance-based approaches</th>
<th>Performance-based approaches</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alterman and Hill (1978), implementation of urban land use plans</td>
<td>Damme et al. (1997), improving the performance of local land use plans</td>
</tr>
<tr>
<td>Calkins (1979), the planning monitor</td>
<td>Driessen (1997), implementing situations in rural land development</td>
</tr>
<tr>
<td>Talen (1996a), methods to evaluate the implementation success of plans</td>
<td>Lange, Mastop, and Spit (1997), performance of national policies</td>
</tr>
<tr>
<td>Baer (1997), general plan evaluation criteria</td>
<td>Mastop (1997), performance in Dutch spatial planning</td>
</tr>
<tr>
<td>Laurian, Day, Berke, et al. (2004), evaluating plan implementation</td>
<td>Mastop and Needham (1997), performance studies in spatial planning: the state of the art</td>
</tr>
<tr>
<td>Brody and Highfield (2005), testing the implementation of local environmental planning</td>
<td>Needham, Zwanikken, and Faludi (1997), strategies for improving the performance of planning</td>
</tr>
<tr>
<td>Brody et al. (2006a), measuring the adoption of local sprawl-reduction policies</td>
<td>Faludi (2000), the performance of spatial planning</td>
</tr>
<tr>
<td>Chapin, Doyle, and Baker (2008), A parcel-based method for evaluating conformance</td>
<td>Faludi (2006), evaluating plans: the application of the European spatial development perspective</td>
</tr>
</tbody>
</table>
propose the PIE, as a form of conformance-based approach, more suited to planners practice, allegedly because it reflects the planners’ principles and intentions; because performance-based methodologies focus on structural and long-term decisions, associated with high levels of uncertainty and, as such, are not very useful in day-to-day decisions; and finally because in practice decisions tend to deviate from plans without compromising their implementation proposals. Brody and Highfield (2005) argue that the absence of a systematic evaluation on plan implementation is because of four main reasons: disagreements on when should plan results be determined or with what should they be compared, the absence of a consensus on how to measure planning effectiveness, the difficulties of the analysis of planning impacts throughout long periods of time, and finally the debate on the concept of success in planning. Chapin, Doyle, and Baker (2008) focus, as the former authors, on environmental issues, presenting a parcel-based geographic information system (GIS) method for evaluating conformance of local land use planning. This method allows the identification of land use changes at the plot level, determining the amount and location of new residential development inside and outside hurricane hazard zones prior to, and subsequent to, state approval of local comprehensive plans.

Performance-based evaluation follows from defining a plan as a decision framework (Alexander 2006a). The plan performance expresses its usefulness in filling this role. It is important to understand if, in what conditions, and how the plan was consulted for subsequent decisions. What happens to the plan is the key for its assessment (Faludi 2000). Based on the work of Fudge and Barrett (1981), who highlighted the differences between conformance and performance, the Dutch school of planning evaluation (Driessen 1997; Lange, Mastop, and Spit 1997; Faludi 2000, 2006; Mastop 1997; Mastop and Faludi 1997; Mastop and Needham 1997; Needham, Zwanikken, and Faludi 1997; Damme et al. 1997) has been developing this approach.

After distinguishing between project plans and strategic plans and concluding that the conformance criterion is useful only in the evaluation of the former, Faludi (2000) and Mastop and Faludi (1997) develop the performance criterion to assess strategic plans. Providing a frame of reference for operational decisions, this type of plan does not have to produce direct impacts on the physical development process. On the contrary, the evaluation of this type of plan should correspond to a detailed analysis of the decisions and actions of a number of actors that are supposed to receive the plan messages. Faludi (2006) extends the performance-based approach to the evaluation of the European spatial development perspective (ESDP). Drawing on the distinction between planning as a technical exercise and as a learning process, the author contrasts the concept of application of plan messages to the traditional concept of plan implementation and presents a method to evaluate the success of the former. The purpose of applying ideas in such a document as the ESDP is to provide the professionals involved in spatial planning processes in Europe with better knowledge of their working contexts and of the directions to follow.

In some few cases studies on plan implementation explored the potentialities of an integrated use of both approaches (Alexander and Faludi 1989) and their simultaneous application for comparative purposes (Altes 2006; Berke et al. 2006).

Alexander and Faludi (1989, 127) distinguish three views on the planning process and the associated criteria to evaluate plan quality: planning as a control of the future, implying that plans not implemented indicate failure (Wildavsky 1973); planning as a process of decision making under conditions of uncertainty, when implementation ceases to be a criterion of success and it becomes difficult to provide rigorous criteria of the quality of a given plan (Faludi 1987); and an intermediate view that still considers implementation important but assumes that, as long as outcomes are beneficial, departures from plans may be considered acceptable (Alexander 1981). The policy-plan/programme-implementation-process (PPIP) methodology, presented by Alexander and Faludi (1989), integrates these views and a number of elements that are not normally considered in an inclusive way, such as plan making, operational decisions, implementation, and the impacts of the plan.

Altes (2006) compares the conformance-based and the performance-based approaches in a case study of the Dutch national urban concentration policies. An application of the former concept reveals that the urban containment policies conform well to the plan. Nevertheless, in the context of the current stagnation in housing production, these policies have not been able to improve the decision-making process. In this sense, the author argues that plans with high conformance can have bad performances.

In the same way, Berke et al. (2006, 581) explore and compare these conceptions of success in planning. These authors sustain that plan implementation in New Zealand is weak. If implementation is defined in terms of conformance, plans and planners have an important influence on the implementation success, but if it is defined in terms of performance, plans and planners are less influential.

**Plan concept and evaluation methodology.** In the literature on planning and evaluation, there are a number of solid contributions to the relationships between planning models and their specific evaluation methods (Alexander 1998b; Alexander and Faludi 1989; Baer 1997; Khakee 1998; Voogd 1997). Despite its reduced expression in the literature, the analysis of evaluation practice in real contexts provides some indications that confirm the importance of these relationships (Alexander 1998a; Khakee 2003; Seasons 2003b). As a practical illustration, Voogd (1997) verifies the rejection of rational assessment methods in some exceptional planning arenas in Holland, where participatory and interactive approaches seem to be preferred.
One of the main challenges of evaluation lies in the lack of a single approach, valid for every situation (Rossi, Freeman, and Lipsey 1999). The literature supports the idea that each evaluation situation possesses a number of specific characteristics that should shape the evaluation methodology. The evaluator should structure his or her methodology according to the specific nature of the situation and should not follow, in a rigid way, a number of standardized procedures. In this sense he or she should have a strong knowledge of the strengths and weaknesses of each available method.

**The evaluation questions, the criteria, and the indicators.** Rossi, Freeman, and Lipsey (1999, 78) define “evaluation questions” as a set of questions—developed by the evaluator, the decision maker, and the main stakeholders—that identify the issues the evaluation will investigate and that are stated in terms such that they can be answered in a way useful to stakeholders using methods available to the evaluator. The formulation of the evaluation questions is probably the most important aspect in the design of an evaluation methodology (European Commission 1999; Rossi, Freeman, and Lipsey 1999).

Evaluation criteria are strongly linked with evaluation questions. Despite the existence of several articles on evaluation criteria (Alexander 2002; Alexander and Faludi 1989; Baer 1997; Berke et al. 2006), there is a generalized view that planning, as a profession, has not developed the necessary criteria to assess the quality of its products and processes. A consensual position in the debate is that the plan concept—or the design criteria—should provide the criteria for the plan assessment. In the case of an internal evaluation, one of the necessary professional skills should be the ability to formulate evaluation criteria as well as to prepare goals and objectives for a plan (Baer 1997). Alexander (2002, 192) believes that professional planners are not the only ones needing substantive plan evaluation criteria, but also national planning systems to evaluate plans in the course of their review, dispute, or approval. This would complement formal evaluation, usually based on conformity with procedural norms.

An indicator produces quantified information to help actors of public interventions to communicate, negotiate, or make decisions (European Commission 1999). The literature on indicators is vast. Much of the research emerged in the 1960s and 1970s and was expanded in the 1980s and 1990s with the emergence of the concept of sustainable development (Seasons 2003a, 2003b). Traditional indicator categories include economic indicators, social indicators, and environmental indicators. These traditional categories were usually developed and applied in isolation, but in the 1980s this panorama changed with the emergence of integrative approaches on themes such as sustainability, healthy cities, and quality of life. A third set—performance indicators—has its origins in performance measurement and management systems (Hoernig and Seasons 2004).

The presentation of the evaluation results and their use by decision makers. In the end of the evaluation exercise, results have to be presented to decision makers and to stakeholders. At this stage, the tension between communication requirements and technical knowledge can become a critical issue. Most evaluation methodologies rely on a technical sophistication and an advanced scientific knowledge that may not be readily understandable by the decision makers and the stakeholders. The evaluator must decide between maintaining the complexity and the technical detail and introducing simplifications to facilitate the communication process. This communication should work in both directions. The evaluator needs to understand what kind of information is the most significant to each of the participants in the evaluation process. These are challenges that many evaluations fail to meet, even when their methods and procedures represent the best practices to date (Alexander 2006d, 273).

After the presentation of the results, the value of the evaluation exercise can be judged by the utility of the conclusions and recommendations. This topic has been the subject of research work, particularly since the 1990s (Rossi, Freeman, and Lipsey 1999), although not so much in the field of planning evaluation but rather in the field of program evaluation. Despite some slight differences of opinion among authors, the literature on program evaluation seems to agree on the idea that only in some few cases are evaluation results directly used to influence the contents of the programs under assessment or the contents of subsequent programs. In most cases the influence of the evaluation results is, at most, of an indirect nature and rather slow to emerge. Weiss (1999) adopts a decision-maker perspective to justify this fact and argues that the search for the best or the wisest program or policy might be only one of the reasons that leads decision makers to promote an evaluation. But even when this is not their primary objective, decision makers are often subtly influenced by evaluation, in what she calls a process of enlightenment (Weiss 1999).

Several authors have been debating how to enhance the use of evaluation results. Ho (1999, 2003) proposes the refinement of a program theory close to the model developed by Pawson and Tilley (1997). Patton (2002) argues that evaluators and stakeholders should, together and at an early stage, define the purpose of the evaluation and the different ways to use the information that will be made available by the exercise. Rossi, Freeman, and Lipsey (1999, 436) propose five guidelines for maximizing the use of evaluation results: (1) evaluators must understand the cognitive styles of decision makers, (2) evaluation results must be timely and readily available when needed, (3) evaluations must respect stakeholders’ program commitments, (4) the subsequent use and dissemination plans should be part of the evaluation design, and finally (5) evaluation should include an assessment of future utilization. Knaap (2004) argues that the actual use of the evaluation results can be enhanced by
assuming, as a starting point for the whole exercise, the stated assumptions and objectives of the plan or policy under analysis.

1.3. Planning Evaluation Practice

Generally speaking, there seems to be a gap between evaluation theory and practice (Khakee 2003). So far there is not an exact notion of the nature and extent of this gap because there are no systematic surveys on evaluation practice in local planning departments or planning agencies (Alexander 2006c). Nevertheless, a reduced number of studies provide a broad picture of the existing situation in some countries in Europe (Carmona and Sieh 2005; Khakee 2003; Lichfield and Prat 1998; Voogd 1997), in the Middle East (Alexander 1998a), and in North America (Seasons 2003b). These studies are a valuable help to understand the main differences between evaluation research and practice, the implications resulting from this gap, the factors that contribute to the application of evaluation in planning practice, and the possibilities of linking theory and practice more effectively. Five evaluation methodologies, largely with an ongoing or an ex post dimension, are subsequently presented: the PPIP (Alexander and Faludi 1989), the Means for Evaluating Actions of a Structural Nature (MEANS; European Commission 1999), the PIE (Laurian, Day, Backhurst, et al. 2004; Laurian, Day, Berke, et al. 2004; Berke et al. 2006), the methods proposed by Norton (2005a, 2005b, 2005c), and Brody and Highfield (2005), Brody, Highfield, and Thornton (2006), and Brody, Carrasco, and Highfield (2006). These methodologies can be seen as alternatives or as complements to the ex ante evaluation methods referred to before.

The relationships between evaluation theory and practice. Carmona and Sieh (2005) analyze recent innovations in performance assessment in more than eighty local planning authorities in England, reviewing the objectives and mechanisms of performance assessment as well as the assessment drivers and inhibitors. These authors report a fragmented picture nationwide, but they highlight a number of promising initiatives that can lead to a more holistic assessment framework. Khakee (2003) presents a set of studies prepared by other authors on the evaluation practice in Sweden. Khakee highlights the dominance of aggregated, quantitative, and rational methods in evaluation practice, contrasting these with recent trends in evaluation theory and research. Under a more general and optimistic perspective, Lichfield and Prat (1998) characterize the evaluation practice in England, pointing out a stronger linkage between the three evaluation stages referred to before and a better integration of evaluation in the planning system. Voogd (1997) analyzes a number of multicriteria evaluation exercises developed in Holland during the past decades. This author sustains that, in planning situations with a clear hierarchic power structure, methods are more often used on a regular basis, while in more flexible situations evaluation methods are somehow disguisedly used in market analysis and GIS studies. Alexander (1998a) describes a number of multicriteria evaluation exercises in Israel, with a particular emphasis on the construction of the Trans-Israel Highway. The author highlights the idea of a gap between the prescriptions of the normative theory and the then current practice characterized by the lack of transparency and by the inconsistency between scientific objectivity and decision processes.

Seasons (2003b) analyzes the evaluation practice in fourteen planning departments in Ontario (Canada), exploring the factors that contribute to or, inversely, that inhibit the development of this activity. Organizational culture is one of these factors. It is related to the attitudes of staff toward and the level of support to planning evaluation demonstrated by senior managers and politicians. Different attitudes and behaviors explain why some planning departments are more receptive than others. The author emphasizes that learning organizations embrace evaluation as a means of enhancing the planning process. Opposite these are organizations that are averse to change and avoid criticism and tend to regard evaluation with suspicion and hostility.

The availability of resources is another important factor. In the majority of municipalities, planning staff resources are concentrated in the review and facilitation of development proposals (Seasons 2003b, 433). Because of all of the constraints that evaluation methods have to overcome, they should be simple, easy to understand, and easy to become operational. Finally, Seasons (2003b) verifies that most of the departments prefer quantitative methods and indicators complemented, in a minority of cases, with qualitative elements.

After analyzing the major differences between evaluation practice and research as well as the main factors that influence the use of evaluation in the planning process, it is important to understand how research and practice can be brought together and which difficulties need to be overcome. Assuming that professional practice will not adopt—at least in a short period of time—the theoretical positions he stands for, Khakee (2003, 349) argues that evaluation agencies in the public sector need an open framework for evaluation. This means that those commissioning evaluation should not exercise complete control over what questions evaluations should pursue, how information should be collected and interpreted, and to whom the findings should be disseminated. This Swedish author rejects the idea of a value-free evaluation as well as the excessive use of quantitative assessments and an uncritical commitment to the scientific paradigm.

Bridging the gap in the other direction, Alexander (2005a, 2005b, 2006c, 2006d) focuses on institutional design. This corresponds to the devising and application of rules, procedures, and organizational structures that will shape behaviors and actions so as to agree with held values, achieve desired objectives, or execute given tasks (Alexander 2005b, 213). This author argues that before discussing the methodological questions, the objectives, and the criteria, evaluation should be seen as a problem of institutional design. As such, less
attention should be paid to criticizing, modifying, and transforming the wealth of existing sophisticated methods and more to developing a useful model of contingent application. Talen (1996a, 79) argues that the frequent shifts in the tool kits of theories and ideologies handed to practicing planners have little hope of solidifying unless they are able to address the challenges of and steadily blend with current planning practices.

The application of evaluation methodologies. The PPPI model was presented by Alexander and Faludi (1989) at the end of the 1980s as a framework to assess the implementation of plans and policies. Notwithstanding the influence of this model in the planning literature, from the early 1990s to date, it has never been used in practice, or reported on, to the best of our knowledge. The PPPI model combines three planning views and the corresponding evaluation approaches, which are regarded as truly complementary. The model lists criteria—conformity, rational process, optimality ex ante, optimality ex post, utilization—in a programmed sequence of questions to be applied to the policy, plan, program, or planning process under consideration as well as to their outcomes (Alexander and Faludi 1989, 134). Depending on the responses to this sequence of criteria, evaluation can be classified as positive, neutral, or negative.

MEANS can be distinguished from the methodologies presented here on two levels, conception and use. At the end of the 1990s, in the context of the allocation of structural funds, the European Commission presented MEANS as a comprehensive methodological guidance to the evaluation of socioeconomic programs. MEANS is based on four main criteria—relevance, efficiency, effectiveness, utility—and three complementary criteria—clarity of the objectives, internal coherence of the objectives (within the program under analysis), and external coherence between the objectives of the program and other relevant public policies. MEANS establishes a typology of indicators in relation to the processing of information, the comparability of information, the scope of information, the stages of completion of the program, the evaluation criteria, and the mode of quantification and use of the information.

PIE was presented by a group of American and New Zealand researchers at the beginning of this decade. PIE relies on the analysis of plans and permits to provide a rigorous, quantitative, and systematic way of assessing the degree to which land use plans are implemented (Laurian, Day, Berke, et al. 2004, 471). Laurian, Day, Berke, et al. (2004, 472) define plan implementation as the degree to which plan policies are implemented through the application of specified development techniques in planning practice. These authors measure two aspects of plan implementation, breadth and depth. PIE has been applied to six New Zealand plans and to almost four hundred land development permits and has focused on storm water and urban amenity management. Undertaking an evaluation with the PIE approach requires the identification of one or several issues of specific interest, the selection of plans or relevant parts of plans, the selection of land use permits, the evaluation of the linkages between plan policies and permits, and the calculation of implementation indicators (Laurian, Day, Berke, et al. 2004, 472). To assess the interaction between plan and permits, it is necessary to identify the techniques (e.g., protection of natural features, development controls) that were used to address the plan proposals. These authors verified that plan implementation was generally low and that the final results tended to have better scores in implementation breadth than in implementation depth. Also, some variations in the overall scores were recorded according to the selection of cities and themes in particular storm water and urban amenity management.

Norton (2005b) describes the application of an evaluation methodology in a set of forty local plans along the coastal zone of North Carolina during the mid-1990s. This methodology is based on six criteria. Three are process related—local elected officials’ commitment to planning, overall plan quality, plan use—and three are substance related—local elected officials’ policy trade-off preference, plan policy emphasis, plan use emphasis. In this study, a high-quality plan is expected to demonstrate a strong factual basis, to provide clearly articulated goals, to employ a land suitability analysis that clearly identifies natural and built environment opportunities and constraints for development, to establish policies that are consistent with the analysis, to satisfy different kinds of consistency, to facilitate meaningful ongoing public participation, to designate implementation responsibilities, and to incorporate monitoring and evaluation procedures (Norton 2005b, 59). Taken altogether, the plans under evaluation proved to be weak from an analytical and substantive point of view, making use of the undertaken analysis in a very limited way and establishing policies for land use classification and development with few references to the necessary constraints to protect natural resources.

Brody, Highfield, and Thornton (2006) examined the spatial pattern of wetland development permits in Florida, verifying its conformance with the proposals of the local plans. Research intended to answer three main questions: (1) how and where wetlands have been developed over a ten-year period, (2) if wetland permits are clustered in areas designated for high-density development (conformity) or if they significantly deviate from the plan’s original spatial designation (nonconformity), and (3) if the quality and content of the original plan relate to its degree of implementation (Brody and Highfield 2005, 160). The authors tested the level of implementation by using GIS to map all permits around wetland areas as indicators of the development subsequent to plan adoption. In this sense, they measured the degree to which development matches the land use configuration prescribed in the original plan around these sensitive areas (Brody and Highfield 2005, 161). The majority of the identified clusters of permits were in conformance with the proposals of the plan. However, a high degree of permit non-conformity was identified in the clusters with significant...
deviations to the local plan (more than 15 percent of the permits). In another article, Brody, Carrasco, and Highfield (2006) used the same methodology and the same case study with success to analyze the effective influence of five sprawl-reduction planning policies in local plans.

2. The Physical Dimension of Planning

After the literature review on planning evaluation presented in the first part of the article, the state of the art on the morphological dimension of planning is described in the second part, revealing the authors’ concern on placing urban form issues alongside the main themes on planning and planning evaluation debates. These two sections inform the proposal of a set of principles for planning evaluation. In particular, this second section (in complement to the former debate on the timing of evaluation and the conceptions of success in plan implementation) will be most relevant to support the fourth and the sixth principles.

2.1. The Plan

Although many cities, since ancient times, have been designed and built according to plans, until the end of the nineteenth century these documents resulted from an empirical activity and not from an integrated and multidisciplinary view of the territory. The modern activity of urban and regional planning emerged only at the turn of the twentieth century, mainly because of concerns over public health and the protection of property values. The main element in this initial stage of planning was the plan, establishing a blueprint vision of the territory and the associated land use regulations to guide development according to that same vision.

The planning model corresponding to this vision was the well-known survey-analysis-plan of Geddes. In addition to Geddes’s main contributions—a structured method of work and a new city concept emphasizing the close ties between the city and the surrounding region—a number of visions of the city steadily emerged in the transition to and in the beginning of the twentieth century. These contributions had a major influence on today’s planning practice. We are talking about the well-known contributions of the Anglo-Saxon tradition, such as Howard’s garden-city theory, Perry’s concept of neighborhood unity, and Wright’s broadacre city. In the European Continent tradition, we point out Soria y Mata’s linear city, Garnier’s industrial city, and Le Corbusier’s radiant city as some of the most influential ideas of that same early period.

2.2. The Planning Process

At the turn of the second half of the twentieth century, planning theory suffered a number of changes that led to a shift of emphasis from plan to process. The rational planning model proposed by the Chicago School (Meyerson and Banfield 1955) and the systems approach (Chadwick 1966; McLoughlin 1969) are fundamentally different from the survey-analysis-plan model that had so much influence on planning practice during the first half of the twentieth century. Faludi (1987) distinguishes the Geddes approach from these contributions because of the fact that the former emphasized research before plan making but did not rendered explicit the crucial step of translating knowledge into action.

Paradoxically, when the rational planning model assumed a hegemonic position in the planning literature, it produced the very conditions for its own crisis. To a large extent, this was because of the increasing separation between two realities, the academic world, more and more obsessed with planning theory, and the real world of planning departments and agencies, focused on daily planning tasks and development control routines (P. Hall 1988). The criticisms of the then emerging models, from transactive planning to communicative planning, focused on the value-free character of the rational model, the erroneous and dangerous notion of a homogeneous society, and the absence of public involvement and participation opportunities.

2.3. The Plan, the Planning Process, and the Results

In the mid-1980s, when the rational-comprehensive model was under criticism, the topic of urban form started to steadily emerge, once again, in the planning debate. With some differences in their proposals, Beauregard (1990), Fainstein (2000, 2005), and Neuman (2005) seem to converge on the idea of formulating a planning theory based on the city. Sternberg (2000) and Talen and Ellis (2002)—as Lynch (1981) did two decades before—argue for a theory of the good urban form. Talen and Ellis (2002) and, in a way, Fainstein (2000) sustain that this theory, corresponding to a normative kind of planning, should be positioned side by side with other planning theories. Talen and Ellis (2002) share with Mandelbaum, Mazza, and Burchell (1996) the view that the dominant theoretical frameworks enable the understanding of the planning process, providing a context for decision making and some important but nonspatial normative contents, but they do not offer any reference framework to planners when the time comes to plan and design specific urban spaces.

Throughout the past decades, three aspects have contributed to this renewed interest on urban form, the debate on sustainable development and the compact city, the emergence of new planning trends—particularly the new urbanism—and the influence of other disciplines such as urban design and urban morphology. Because of the nature and scope of this article and the quality of the results already achieved, the third of these aspects deserves a closer look.

An important contribution to the resurgence of urban form in planning debate was given, in the 1990s, by the concept of sustainable development because of its strong link to the form of the city (Breheny 1992; T. Hall and Doe 2000; Jenks,
Burton, and Williams 1996). Breheny (1992) focuses on the spatial dimension of sustainability. He believes that as cities are, simultaneously, the main consumers of natural resources and the major producers of residues, a good design and management of these cities can contribute to the resolution of global environmental problems. More recently, sustainable development has framed the discussion on city models such as the compact city model, stimulated by multiple initiatives of the European Commission. In the meantime, and following some criticisms of the compact model, other concepts and ideas have been proposed as alternatives to the compact city model, namely, the decentralized concentration concept (Banister 1992; Owens 1992) and the growth regions concept (Valk and Faludi 1992) or the acceptance of sprawl as inevitable and indeed with alleged positive aspects (Gordon and Richardson 1997). Other authors, such as Breheny (1992, 1996), sustain a compromise between both tendencies. Two decades after the beginning of the debate on sustainable development, it is rather consensual that it should be considered as a political goal and not as a scientific notion (Haberl and Schandl 1999). Therefore, the discussion in the beginning of the twenty-first century focuses on a more real and measurable question, an urban environment with low levels of carbon. In this context, Haberl and Schandl (1999) identify two research lines, the exploration of the concept of urban metabolism (mainly developed by the industrial ecology school, based on the initial conceptualization of Wolman 1969) and the comprehensive study of land use and land occupation changes.

As it was mentioned before, the resurgence of urban form as one of the main themes of the current planning debate is also from the emergence of new planning trends, particularly the new urbanism. Its importance has been highlighted by several authors and on several grounds. Fainstein (2000) presents the new urbanism as a theoretical proposal alongside communicative planning and the just city theories. Ellis (2002) portrays new urbanism as an ambitious alternative to the conventional practice of land use development in America, characterized by low densities and car dependence (Ellis 2002). Talen and Ellis see the new urbanism as a proposal for a theory of the good urban form (Talen and Ellis 2002), embracing a vast array of ideas and contributions that have been developed since the nineteenth century (Talen 2005), not only at the urban scale but also at the regional scale (Talen 2008).

A third contribution to reposition urban form in the center of the debate is provided by urban design and by urban morphology. Urban design has been acquiring widespread popularity, most evident in professional journals, government Web sites, academic debates, and popular media (Madanipour 2006, 173). Beside moving from the margins to the center of the debate, the dimensions and the scope of urban design have also been widening (Carmona et al. 2003; Madanipour 2006). Throughout the past three decades, and particularly in England, the role of design guides as supplementary planning guidance has been rather widespread. The development of these guides has followed the pioneer publication of the Design Guide for Residential Areas prepared by the Essex County Council back in 1973. This guide highlighted the needs to enhance the design of new residential areas and to make the design responsive to the locality, less wasteful of space, and more attractive to living environments (Essex County Council 1973). T. Hall (2005, 2007) presents his experience as chairman of the Planning Committee of Chelmsford Borough Council. After years of production of poorly designed built forms and the loss of some historic buildings, a new political administration started in the mid-1990s, the process of achieving higher standards of quality and sustainability in the built environment. The main components of this approach were an investment in staff, the publication of a local design policy (which included the adoption of the revised Essex Design Guide), the definition of a long-term vision, and finally greater cooperation among the different professions within the Borough Council and the agencies involved in the process of urban development.

In 1996, the first planning brief was prepared for a major housing site within the 1991–2001 local plan. Though the first design briefs were rather general, they already included some important urban design principles. T. Hall (2005) states that, over time, there was a steady increase in the amount and degree of prescription in design control policies. As several tools of design control became more detailed, clearer, and more purposeful, so the quality of the physical results improved.

One of the main research themes in urban morphology has been the articulation of urban form studies and planning, particularly through the application of morphological concepts in planning practice. Kropf (2001) and Whitehand and Morton (2003, 2004) explore the application of the concept of fringe belt, corresponding to a zone of mostly extensive land use that comes to be on the edge of an urban area during a hiatus in outward residential growth. The Stratford-upon-Avon District Design Guide coordinated by Kropf (2001) and adopted as supplementary planning guidance in 2000 proposes the fringe belt as a fundamental element of the urban structure and of the historical and geographical development of the city.

Kropf (1997, 2001) has developed other key concepts for bridging morphology and planning type, urban tissue, and levels of resolution. Kropf (1997) sustains that the concept of type can help solve some of the main problems associated with functional zoning. He highlights three principles of the Italian school of urban morphology that can be used to build a form-based zoning: existing forms are a result of learning and a record of past experiences in accommodating human activities and needs, built forms and human activities are interrelated but the relation is not fixed, and the structure and character of a town result from both continuity and change at
various levels. These principles lead to three working assumptions: zoning regulations should take local and regional forms as the starting point for prescription, allow for mixed uses, and permit both continuity and change. In the *Stratford-upon-Avon District Design Guide* and in the *Plans d’Occupation des Sols* (POS) of Asnières-sur-Oise and of Mennecy, Kropf uses a synthetic conception of the urban tissue as an organic whole, whose form can be described at distinct levels of resolution: streets and blocks, plots, buildings, rooms or spaces, structures, and materials. These different urban elements are interrelated in a hierarchy. Smaller scale elements combine to form larger scale elements that in turn are parts of still larger elements. Using the hierarchy as a framework, it is possible to systematically define tissues at different levels of specificity by describing the constituent elements stepwise through the levels of resolution (Kropf 1997, 131). Three specific characteristics are used to describe each element: position, outline (shape, size, and proportions), and arrangement (type of component parts, number of parts, and relative positions).

Based on the levels of resolution, McGlynn and Samuels (2000) present the concept of funnel. The representation of the variety of urban forms at different levels of resolution is given by the image of two funnels. In the first, corresponding to traditional towns, there is a wide diversity at the top of the funnel (districts, streets, plots) and a progressively reducing diversity toward the bottom. In the second, corresponding to modern cities, the funnel is reversed. After verifying that, in recent years, house builders and promoters have been improving the design quality in the lower levels of the funnel, the British authors focus on the issue of spatial continuity at the higher levels. Within his professional activity (after the experiences of Asnières-sur-Oise and Mennecy), Samuels has been developing methods that are easier to use and that consume fewer resources than the methods presented above. The POS for St. Gervais-les-Bains illustrates this development (for a detailed presentation of this plan, see Pattacini 2001 and Samuels 1999).

Four years after the preparation of its design guide, the Stratford-upon-Avon authority promoted a study on eight residential areas (Larkham et al. 2005) and subsequently adopted it as supplementary planning guidance. Based on the concept of morphological region, the study defined and sustained a classification scheme and the proposal of two of these as conservation areas. Furthermore, it provides guidance on streets, public spaces, plot patterns, building patterns, architectural and historical qualities of buildings, predominant local building materials, singular building, land uses, and vegetation.

### 3. Principles for Planning Evaluation

After the literature review in the first and second parts of the article, this last section intends to synthesize a number of principles that can contribute to defining and structuring planning evaluation methodologies (also see Table 3). The boundaries among these seven principles should be seen as a tool to enable more effective analysis and not as a denial of existing interactions.

First, the evaluation of planning is a complex and difficult but also a rather necessary exercise. This need is supported on the grounded belief that evaluation can contribute to a better planning practice. The main arguments are fourfold: (1) evaluation legitimates planning before citizens, providing sustained appraisals on planning products, procedures, and results throughout the whole planning process; (2) it helps politicians and planners in complex processes of decision making; (3) it tracks the course of planning proposals, promoting an effective planning dynamic, in which suggestions for changes or reviews in planning products and processes are supported by the results of evaluation exercises; and finally (4) it enables the construction of a planning practice based on a continuous learning process. The main challenge to start implementing evaluation mechanisms in planning practice is to make evident before politicians and planners the advantages of evaluation. This can be done through the promotion of an evaluation culture within the institutional structures that undertake planning activities (e.g., planning departments or similar agencies). While some structures might have the required characteristics to undertake this process in the short term, others may not. The identification and dissemination of real cases where the evaluation practice is contributing to a better planning process should lead to the progressive adoption of evaluation in other cases. The following principles address a number of fundamental issues on the evaluation of planning practice.

Second, the design of an assessment methodology must be clearly linked with evaluation theory. On one hand, the methodology will be improved by inputs from this body of knowledge. On the other hand, this linkage will contribute to bridging the gap between evaluation theory and practice, a rather relevant objective within this activity. Despite the lack of an effective characterization of this gap, mainly because of the reduced number of research works on the evaluation practice in local planning departments, this seems to be a problem to include in the agenda, as is the need to start building bridges in both directions. The development of two specific procedures can contribute to the establishment of the linkage. Each methodological proposal should contain a view of evaluation, reflecting its position on a number of fundamental issues in the current debate. Furthermore, it is necessary to build a continuous process with effective mechanisms enabling the permanent exchange of data between theory and practice. In other words, in spite of designing evaluation methods more and more sophisticated, academics should work on the articulation between theoretical and methodological developments, on the practical application of their methodologies—if possible, integrated into real planning processes—on the analysis.
Table 3. Principles for Planning Evaluation

1. Planning practice should be evaluated
   Why? Evaluation can contribute to a better planning practice
   How? The advantages of planning evaluation must be clear for politicians and planners
   Planning evaluation’s theory and practice must be continuously developed (see the next principles)

2. The design of an assessment methodology must be clearly linked with planning evaluation theory
   Why? The methodology will be improved by theory’s inputs
   How? The methodology will help bridging the gap between evaluation theory and practice
   Each methodological proposal should reflect a view of evaluation
   Mechanisms to enable exchange of information between theory and practice should be built

3. The evaluation methodology should suit the object under appraisal
   Why? An object can only be fairly assessed if the methodology is adequate for that purpose
   How? Each planning practice (the object) is framed by a number of specificities (territory, legal system, planning culture, etc.) that contrast with the generic character of the model
   Why? It should be possible to introduce changes in the methodology based on the object

4. The planning practice (plan, process, results, etc.) must be evaluated as a whole
   Why? Because of the evolution line of planning theory, the need to integrate evaluation in all stages of the process, the debate on the two conceptions of success, the definition of evaluation functions
   How? These different dimensions must be comprised within the general criteria of evaluation, the specific criteria, the indicators, the data sources, and the assessment techniques

5. Evaluation and planning processes should be developed together, right from the beginning
   Why? Evaluation and planning are not discrete activities
   How? Defining the planning team, the evaluation team, and the framework for their relationships
   Establishing together: a view of planning and of evaluation, the planning aims and the evaluation purposes, the design criteria and the assessment criteria, and the data sources

6. The evaluation process must have a balanced development in time (ex ante, ongoing, ex post)
   Why? Contrarily to current predominance of ex ante dimension, evaluation must comprise the whole process
   How? Develop a theoretical and methodological comparison with program evaluation
   Use the emergent set of methodologies focused on plan implementation
   (As in fourth principle) the main parts of the methodology must comprise these different dimensions

7. The presentation of evaluation results and the analysis of their use by planning should be valued
   Why? It is necessary that politicians, planners, and citizens be able to understand the evaluation results
   How? Presentation mechanisms must find the balance between communication and technical knowledge
   Establish the terms (and the required mechanisms) by which a successful utilization may be judged

of the use of evaluation results by planners and politicians (to be developed in our seventh principle), and on the contributions of these experiences to the wider planning debate.

Third, the evaluation methodology should suit the object under appraisal. An object can be assessed only if the type of evaluation is adequate for that specific purpose, under the risk of obtaining a set of false or unfair results. Therefore, an evaluation methodology must understand the kind of planning practice that it intends to assess. Each planning practice is mainly framed by the socioeconomic and environmental conditions of the territory, by a national legal system, and by a planning culture. To suit the generic evaluation methodology to the particular planning context, it should be possible to change it based on that specificity. While the generic criteria of the methodology should remain the same, the specific criteria and the evaluation questions they correspond to can be adapted according to the specific context. Table 4 presents a number of current planning approaches focused on the plan, the process (rational, decision centered, and communicative) and the city, and the corresponding evaluation practices. We believe this scheme can be the basis for the construction of a contingent planning approach, where the importance of the different elements is context dependant. Let us assume a hypothetical evaluation of a land use plan. One of the methodology’s generic criteria is plan rationality, and within this criterion, one of the specific criteria is the external coherence between the local plan and other plans prepared for that same territory. We believe that when applying this methodology to cities with a low planning dynamic and before the absence of other plans, adjustments may be introduced (although weighted in the final result) and the articulation of the plan with other instruments (e.g., housing or transport policies) can be considered.

Fourth, the planning practice (plan, process, results) must be evaluated as a whole. Four main reasons justify this statement: (1) the evolution line of planning theory throughout the twentieth century, particularly the way it has been dealing with the urban form issues; (2) the need to integrate...
evaluation in all stages of the planning process; (3) the debate on the two conceptions of success in plan implementation, conformity and performance, in particular in the development of integrated approaches such as the PPIP; and finally (4) the definition of three fundamental functions of evaluation—judgment, learning, and interaction with planning. An evaluation methodology must be able to assess the different dimensions of planning practice: plan making, plan implementation (including development control), and plan review. As such, the general criteria of evaluation should focus on these different dimensions. The general criterion of plan quality (see Table 4) should include specific criteria based on themes such as the articulation between the plan and its theoretical and legal framework, the relevance of its proposals, its internal and external coherence, and the public’s participation in plan making. The general criterion related to process should comprise specific criteria regarding the relationships between politics and planning, the availability of human and financial resources, the promotion of communicative action and interactive practice, and the public participation in plan implementation. Finally, the methodology should include specific criteria related to the achievement of physical results on the ground and the influence of planning in the urban development process. The selection of indicators, data sources, and techniques should also reflect this comprehensive nature associated to the definition of criteria.

Fifth, the evaluation process and the planning process should be developed together right from the beginning. Evaluation and planning should not be seen as two separate activities. Evaluation and planning practices are cyclic processes, and a number of interaction moments between both should exist. Right from the beginning, the evaluation process should be simultaneously structured alongside, and in articulation with, the planning process to be able to provide a set of contributions that can be used by planning in due time. Thus, the first tasks should correspond to the definition of planning and of the evaluation teams and to the establishment of the mechanisms that should frame their relationships. A number of aspects related to this issue have already been discussed in the first part of the article. Both teams should work together on the definition of a common view of planning and of evaluation as well as on the fundamental aims of both activities, on the design and evaluation criteria, and on the data sources to be used throughout these practices.

Sixth, the evaluation process must have a balanced development over time, involving the ex ante, ongoing, and ex post dimensions. Nowadays, ongoing and ex post evaluations in planning have a rather reduced expression when compared to
ex ante evaluations of alternative plans. This is similar to planning practice, where plan making tends to be overestimated in comparison to plan implementation and development control. On the contrary, in the social science tradition of program evaluation, the ex ante appraisal is underestimated relative to the other dimensions. A theoretical and a methodological comparison between planning evaluation and program evaluation can be a first step for the construction of a balanced process. Furthermore, a number of methodologies, specifically developed to assess the success of plan implementation, have been identified in this article and should be put into practice. Finally, the main parts of the evaluation methodology—the definition of general and specific criteria, indicators, data sources, and assessment techniques—must comprise these different time dimensions, in complement to the different content dimensions presented in the fourth principle.

Seventh, the presentation of the evaluation results and the analysis of their subsequent utilization in planning practice are two important stages of the evaluation process that should not be devaluated. In the end of the evaluation exercise, results must be presented to different audiences—politicians, planners, citizens—in an understandable way, reinforcing the main arguments and the fundamental recommendations for a responsive planning practice. The key challenge is to find the right balance between communication (and a practice focused on the quality of interactions and on consensus building) and technical knowledge (and a practice that emphasizes decision taking and aspires to optimal action). An effective communication and the subsequent use of evaluation results may be enhanced if potential users are informed of the strengths and weaknesses of these processes and to what extent definitive results may be expected. The value of an assessment exercise can be judged by the utility, direct or indirect, of its results to planning. Thus, and similar to program evaluation exercises, planning evaluation methodologies should include an analysis of how planning practice is actually using the assessment results. Evaluators, planners, and politicians should not only share the understanding of evaluation purposes but also agree on the criteria by which evaluation may be judged and build the necessary mechanisms to undertake this analysis.

4. Conclusions and Future Research

The literature review covered a wide range of topics, namely, the evolution of evaluation theory and methods, the contemporary debate on planning evaluation, the practice of planning evaluation, and the importance of urban form within planning throughout the past century. Subsequently, seven principles were elected to structure and guide the evaluation process in urban planning: (1) planning practice should be evaluated, (2) the design of a methodology to assess this practice must be clearly linked with planning evaluation theory, (3) the evaluation methodology should suit the object under appraisal, (4) the planning practice must be evaluated as a whole, (5) evaluation and planning processes should be developed together, right from the beginning, (6) the evaluation methodology must have a balanced development over time, and finally (7) the presentation of evaluation results and the analysis of their use should be valued.

Throughout this article a number of gaps needing further research have been identified, particularly between the theoretical proposals for a good city form and their practical application, between planning evaluation theory and practice, and between planning evaluation and program evaluation. The second part of this article focused on the physical dimension of planning. Despite the presentation of a set of plans incorporating morphological concepts and of good practices in urban design, we must highlight the need for further empirical work on the practical implementation of theoretical concepts such as the compact city. Much of the debate on this model is still more in the realm of beliefs than in theoretical arguments confirmed by practice.

The second gap leads us to suggest the realization of an international study on comparative evaluation practices in different planning departments. Despite all the different arenas of contemporary debate, it is surprising how little we know about planning practice within these structures. Empirical work is needed to understand if evaluation exercises do exist, if practitioners are incorporating the new theoretical and methodological developments, if evaluation exercises are isolated acts or continuous processes, what the evaluation focus is, what the most common criteria and evaluation questions are, and particularly if evaluation results are in fact influencing the planning practice.

Finally, as far as the links with program evaluation are concerned, we believe that future research work should focus on three main issues: the definition of evaluation questions and criteria, the presentation of evaluation results and its subsequent use in planning practice, and the exploration of the ongoing and ex post dimensions. The exploration of these dimensions requires evaluation methodologies focused on implementation. Based on the seven general principles presented above, the authors are currently designing an ongoing evaluation methodology, supported by three criteria—rationality, conformance, and performance—that will be applied to the land use plans currently in force in Lisbon and Oporto. This evaluation exercise will be carried out by a team independent from the planning teams involved in the preparation of these plans. We believe that the experience gained with these two different cases will provide a test of our methodology and useful suggestions for improving current planning practices in each of these cities.

Acknowledgments

The authors would like to thank the three anonymous referees for their helpful comments and suggestions.
Declaration of Conflicting Interests

The authors declared no potential conflicts of interests with respect to the authorship and/or publication of this article.

Financial Disclosure/Funding

This research was supported by the grant SFRH/BD/17523/2004 from the Portuguese Foundation for Science and Technology.

References

Alexander, Ernest R. 2000. Rationality revisited: Planning para-

Alexander, Ernest R. 2002. Planning rights: Toward normative cri-

Alexander, Ernest R. 2005b. Institutional transformation and plan-

Alexander, Ernest R. 2006b. Evaluations and rationalities: Reason-

Alexander, Ernest R. 2006c. Evolution and status: Where is plan-


Lyons, John. 2007. Planning rights revisited: Planning para-

Lyons, John. 2008. Planning rights revisited: Planning para-

Lyons, John. 2009. Planning rights revisited: Planning para-

Lyons, John. 2010. Planning rights revisited: Planning para-

Lyons, John. 2011. Planning rights revisited: Planning para-

Lyons, John. 2012. Planning rights revisited: Planning para-

Lyons, John. 2013. Planning rights revisited: Planning para-

Lyons, John. 2014. Planning rights revisited: Planning para-

Lyons, John. 2015. Planning rights revisited: Planning para-

Lyons, John. 2016. Planning rights revisited: Planning para-

Lyons, John. 2017. Planning rights revisited: Planning para-

Lyons, John. 2018. Planning rights revisited: Planning para-

Lyons, John. 2019. Planning rights revisited: Planning para-

Lyons, John. 2020. Planning rights revisited: Planning para-

Lyons, John. 2021. Planning rights revisited: Planning para-

Lyons, John. 2022. Planning rights revisited: Planning para-


**Bios**

Vitor Oliveira is a Post-Doc Researcher at CITTA, whose research has been focusing on a triangle built around three main subjects, Planning, Evaluation and Urban Morphology.

Paulo Pinho is a professor of planning at the Faculty of Engineering, University of Oporto, founder and Director of CITTA and of the MSc in Planning and Design of the Urban Environment.